

World's end or its beginning?

The British are a fitfully impulsive lot. Having let their universities and other research institutions moulder for years, they now propose to engineer a revolution in them.

THE rush of blood to the British government's head about research this past week is not, of course, an accident. Sooner or later, something had to happen. A government elected to a third term in an over-polarized democracy such as Britain's cannot be a stupid government, certainly not one so stupid as to believe that research is unimportant. The neglect of the past eight years must be put down to other things — a decent preoccupation with the condition of the British economy, foreign wars or the threat thereof and too great a respect for a now departed and misguided colleague (Sir Keith Joseph). Now, at last, there is time for ministers' interest to linger on research, and for them to convey to others that they are ready to listen. Whether researchers will be merrier on that account is another matter.

Before laying out real money, readers should consider what is happening elsewhere. The government of France is embarked on another attempt to make France even more competitive (see page 286), that of Australia is on the same bent (*ibid.*). Evidently all governments, while committed as they are to the opinion that they have no money of their own to spend, but only other people's taxes, are zealous in their determination to see that what they spend on apolitical causes such as research is returned to them, preferably with interest. The goal is laudable, only the methods are in dispute.

Some will say that it would have been better if scholars had never taken government money, but had managed their affairs themselves, which is a hopelessly fanciful idea; not merely is government money necessary, but sometimes it is thrust upon people. So the British Advisory Board for the Research Councils (ABRC) has done right (or at least prudently) in hitching its wagon, with its government's, to the star of national prosperity, allowing the caveat that there should be modest freedom for people to satisfy their curiosity. But why, instead of guessing how academic institutions should be categorized, did it not think of advocating a framework for running academic institutions that would give those who belong to them a true sense of responsibility — that of knowing that their irresponsibility could mean that the institution that pays their monthly cheque may not exist? For thirty years, the academic problem in Britain has been recognizing that some academic institutions are different from and even better than others. The arguments about the terms on which guest workers at academically sited interdisciplinary research institutes should be employed are, by comparison, fine tuning.

The robust view is that the scene has been changed, and that it is now high time for the actors also to change their ways. Britain was no longer the world power it had been when it declared (in 1964) the great expansion of its academic system advocated by the report of the Robbins committee on higher education. It over-reached itself, and has since had to pay a high price for its indiscretion. At bottom, ABRC is trying to redress the balance, and intelligently: the proposal that graduate students should be free to decide where they will be taught, and that the government and the research system should pay more attention to the opportunities for international collaboration, makes admirable sense. Even the plan to pay surviving academic researchers the indirect costs of their work out of funds transferred from the

universities's budget, which will quickly distinguish between institutions of "Type T" and the others (see page 280) is sensible, provided that there is some flexibility left in the system. The obvious danger is that nothing can guarantee the chance, for which ABRC pleads, that institutions classified "T" should be enabled (mostly by their own efforts) to swim with the "R"s. ABRC is willing to be flexible (with the help of money hypothetically plundered from the University Grants Committee), but is the government prepared to follow suit?

The government's proposals for the new machinery of research administration are not encouraging. When the House of Lords five years ago asked for a part-time minister in charge of research, it was told that the prime minister (being a scientist) was in charge. Now, having asked a second time more modestly, the House of Lords is told that the prime minister will have a larger staff to help her, and will even meet her much larger advisory council from time to time. How much time the chairman of Rolls-Royce, Sir Frances Tombs, will have to help out is another imponderable, like that of the direction in which his influence would be weighted. The obvious difficulty is that even if the new arrangements work well, and as advertised, they will not meet the clamant need that there should be opportunities to put the case that research has special interests sometimes conflicting with, but not necessarily transcending, those of the British economy. That issue will be one of the first arising under the new regime: does it make sense that Britain should pull out of CERN, the European Organization for Nuclear Research at Geneva, at this stage?

What, in these circumstances, should British researchers do? The best course for the time being is to stick at the bench, to get something done. If the bench happens to be in an institution about to be labelled "T", it is probably too late already to transfer in good order to one labelled "R", although opportunities may arise. Since the institution is in any case embedded in a still-fragile economy beginning to work its way out of difficulties, there is much to be said for helping along the government's recipe of prosperity through research. In that spirit, there is not much room for railing at the disappointments of the past eight years, although it is quite proper to ask why it has taken eight years to discover that something different should be done; that, in any case, is best left to those long since forced out of the enterprise by Sir Keith Joseph's misconceptions of research, intrinsically much more malevolent than those of the present management.

Meanwhile, it is best to keep one's powder dry for the battles that lie ahead. The present British government, which is energetic and successful, is the only government the British have. Its proposals on research, while leaning heavily on schemes being tried elsewhere (how will a British interdisciplinary research centre differ from those being founded, less quickly, by the National Science Foundation in the United States?), may or may not succeed. The university system as the British know it will certainly be transformed, some will say lost. But what will happen if the gamble succeeds, and the British economy also becomes prosperous again? That is when dry powder will be needed. But that battle is a long way ahead. □



Revolution on cards for British academic science

- Pecking-order for universities
- UGC funds transfer to research councils

London

FAR-REACHING changes in the organization, management and support of scientific research in British higher education and in the research councils are advocated this week by the Advisory Board for the Research Councils (ABRC), the body whose chief function is annually to recommend the distribution of available research funds between the five active research councils.

In a document *A strategy for the science base* (HMSO, £3.95), drawn up at the request of the Secretary of State for Education and Science, Mr Kenneth Baker, ABRC advocates that some institutions of higher education should be recognized as primarily teaching institutions, that funds should be transferred from the University Grants Committee to the research councils to help with the payment of research overheads and that there should be founded up to 15 interdisciplinary research centres linked with academic institutions not in the "teaching category".

In passing, the document says that Britain spends "too much" on particle physics, which augurs ill for the impending decision on continued British membership of the European Centre for Nuclear Research at Geneva.

ABRC's document is intended for discussion only at this stage. The government says it will announce its decisions after consultation on the issues raised.

The proposals on categories of universities will be controversial. ABRC defines institutions of "Type R" (for research) as those offering undergraduate and postgraduate teaching and "substantial research activity across the range of fields", estimating that there are about 15 in this category. "Type X" institutions are those whose teaching is less broad and whose research is less broad still; ABRC's estimate that there may be more than 15 of them suggests the inclusion of several polytechnics as well as of institutions such as the Cranfield Institute of Technology.

Institutions of "Type T" (for "teaching") are those that teach undergraduates and MSc students but that do not have "advanced research facilities". Such institutions would be discouraged from making bids to house the interdisciplinary research centres which, ABRC expects, will eventually entail a large part of the research councils' support for research.

The concept of the interdisciplinary research centre, put forward as a means of helping the translation of bright ideas into national wealth, is that they would have a "positively managed coherent programme of work undertaken by a small number of core staff". Most of the research would be carried out by "visiting teams" of researchers who had won research grants from the relevant research council to work in the centre's field of interest; the researchers themselves would be seconded from their home institutions.

ABRC's proposals on the transfer of funds from the University Grants Committee to the research councils (represented by itself) will also be controversial, amounting as they do to a recognition that the "dual support" system (direct costs from the research council, overheads from the university) no longer functions.

The suggestion is that there should be a phased transfer of funds except in the case of major new initiatives and directed programmes of research, such as the successor to the Alvey programme of research in information technology, (where the transfer should be immediate). ABRC says that its recipe will not work without an "explicit and sustained" programme of restructuring, and unless the government provides the funds with which to meet the costs of staff relocation and early retirement as well as those of new posts, new buildings and new equipment.

The thread running through ABRC's argument, that Britain's ingenuity must be bent to industrial prosperity, is explicit: "judgements of potential utility should play a greater part in determining the relative level of effort to be devoted to different fields of basic research..." Nevertheless, "it will be important to maintain some active research capacity across a wide range of basic work.... The government should consider giving genuine protection to agreed levels of curiosity-oriented responsively-supported basic work".

The first formal reaction has come from the Committee of Vice-Chancellors and Principals (CVCP), which welcomes the concept of the research centres, but worries that the new three-tier system will "inhibit change".

At a recent meeting, CVCP condemned the idea of an "explicit hierarchy", claiming that it would engender complacency at the top and depression at the bottom.

Simon Hadlington

New machinery in the pipeline for British science

London

THE newly elected British government is embarking on a thorough reorganization of its administration of science and technology, with Mrs Margaret Thatcher, the Prime Minister, very much in charge. So much emerges from the government's response*, published on Monday, to the latest recommendations of the House of Lords Committee on Science and Technology, themselves published just six months ago.

On machinery, the government intends "collective ministerial consideration, under the Prime Minister's leadership, of science and technology priorities". This is taken to be a reference to a new cabinet committee. There is also to be "an expanded independent advisory body", reporting to the prime minister who "will hold periodic meetings with this body", and which will be called the Advisory Council on Science and Technology (ACOST). The new council will have a wider remit, covering defence as well as civil research, than the Advisory Council for Applied Research and Development, which it will replace.

The government's response refers to the accompanying strategy document from the Advisory Board for the Research Councils (ABRC), saying that the government will solicit comments on the document and then announce its decisions. But on one point, the transfer of funds from the universities' budget to the research councils, the government indirectly supports ABRC in saying that institutions carrying out research contracts "should not subsidize the clients — public or private".

On the theme, running through the response, that the overriding objective is to win "economic and social" benefits, the government says that the chief need is that industry should spend more from its own resources. Nevertheless, the Department of Trade and Industry is reviewing its role in the support of innovation. It believes that the Lords' recommendation that companies should be required to publish accounts of money spent on research and development will be met by impending changes in British accounting standards, is cool on proposals for tax incentives to encourage research and development, but considers that public purchasing has a part to play.

On defence research, the government promises to facilitate arrangements to make fuller civil use of what is done in defence research establishments, and to keep security constraints to a minimum. □

*Civil research and development Cm 185, HMSO, £2.20.

World's ocean drillers plan for decade of technological change

Washington

THE Ocean Drilling Program (ODP) is gradually changing course as more sophisticated technology provides new opportunities for experiments and new ideas about the Earth's evolution demand more sophisticated technology to run bigger scale experiments. So much emerges from a meeting of the world's ocean drillers held in Strasbourg last week to try to define the overall objectives for the ODP in the 1990s.

The Second Conference on Scientific Ocean Drilling (COSOD II) comes a full six years after the first. Day-to-day management of the programme rests with the Texas A&M University, with overall planning provided by the Joint Oceanographic Institutions for Deep Earth Sampling (JOIDES), an international body that represents the US, Canadian, French, British, West German, Japanese and European organizations that support ODP. Only occasionally is the need felt to assemble the world community to try to reach agreement on the larger framework within which individual projects will be planned.

Recommendations will be made early next year after five working groups (for global environmental change, mantle/crust interactions, fluid circulation and global chemical budgets, brittle and ductile transformation of the lithosphere, and evolution and extinction of oceanic biota) have revised their proposals and met once more. But it is already clear that everyone's research needs can be met only if the programme goes in two different directions at once.

Some, particularly those studying global environmental changes and evolution and extinction, need breadth rather than depth. They want to construct global maps embodying variation in the sedimentary record in space, time and depth to tell how ocean, atmosphere and biota respond to change. That need can best be met by a ship capable of going around the world (including even polar regions), and piston coring of ocean sediments.

But other groups, particularly that looking at mantle/crust interactions, want depth rather than breadth. Although sympathetic to the palaeo-oceanographers' desires to "collect a world library of mud", the mantle/crust researchers emphasized at Strasbourg the need to develop the technology for drilling deep holes. The aim is to get a 6-7 km deep hole down to the Mohorovičić discontinuity (Moho) within the next 15 years. Other groups would also like to be able to drill deep in difficult places, have the capacity to stay on one station for months and to re-enter a

hole several years later. The fluid circulation group, for example, would like to have data from inside an active hydrothermal system. The necessary advances



JOIDES Resolution (leased from Sedco/BP).

in drilling technology are difficult, but do not appear impossible.

Although no programme could possibly meet every wish expressed at the meeting, a massive leap forward could be made if there were a second vessel to handle the

less technically demanding needs. The JOIDES Resolution drill ship is very sophisticated and using it for collecting sedimentary cores can be compared to "using a Rolls-Royce when a minibike would do", according to one researcher.

A second ship has usually been thought of as just a dream — but, according to US researchers, French scientists seem convinced the dream can be made real. A budget increase of somewhere between 10 and 50 per cent will be needed, which it is hard to see coming from the United States, already by far the largest contributor. US scientists hope that Europe can provide the necessary push.

Organization is another issue. The present routine, in which the ship voyages around the world visiting a variety of regions, reflects an organizational structure in which five regional panels have considerable say. A sustained effort requiring long periods at one location would be hard to sell. For this reason, and because many feel that there is a need to move from an 'exploration mode' to an 'experimental mode' in which specific hypotheses are tested, there may be a gradual shift towards clearer definition of goals with project managers coordinating drilling on many different legs.

Alun Anderson

French Earth science policy set for radical shake-up

Paris

EARTH sciences research in the Centre Nationale de la Recherche Scientifique (CNRS), France's leading science research body, has been radically shaken up by a decision that all future studies should be concentrated on a limited number of major questions.

This initiative, the brainchild of Claude Allègre, formerly director of the Institut de Physique du Globe (IPG) in Paris and now director of its geochemistry laboratory, aims to move away from the previous CNRS strategy of Actions Thématiques Programmées (ATPs), in which research was focused on specific topics in each discipline, usually for five years. Instead, there will be a simpler administrative procedure with perpetual priorities.

By abandoning the ATP principle, whose themes were beginning to appear arbitrary, Allègre hopes to "aim at more fundamental problems, such as instability or global transfer, and to apply the approach of geophysics and geochemistry to surface phenomena such as erosion and sedimentation".

Two programme areas have been established, each with its own selection committee. The first, entitled Imaging and Structure of the Earth (IST) and chaired by Jean Auboin, professor of structural

geology at the University of Paris VI, will include seismic imaging of the Earth's interior, crustal imaging and satellite-based imaging.

The second programme, chaired by Allègre himself, will deal with the Dynamics and Budgets of the Earth (DBT), and will have ten major headings: ocean budgets (which, for the first two years, will concentrate on the influences of the world's 20 major rivers), mechanisms of erosion, sedimentation and sea-level change, continental growth, global Earth coupling, instability phenomena in the Earth's crust, remagnetization and its importance for palaeomagnetism, rifting and kinetics of geochemical reactions.

As well as restricting the themes that will be supported, Allègre wants the bulk of the research to be tackled by a small number of 'centres of excellence', with the aim of reducing by one-third the number of projects supported and encouraging smaller groups to associate themselves with larger ones.

Allègre stressed that this should not be interpreted as a 'lame duck' policy. "We are talking only about programme money, not running costs or salaries. Laboratories that do not want to take part will still be able to get along."

Peter Coles

Ambitious Japanese proposals for superconductors

Tokyo

JAPAN's Science and Technology Agency and Ministry of International Trade and Industry (MITI) have floated some ambitious proposals to promote research and development of high-temperature superconductors. But whether their plans will be realized will depend on negotiations with the Ministry of Finance and with industry over the coming months.

There is a widely held view in the West that Japan has already begun a coordinated national research effort to conquer the world market in high-temperature superconductors (if and when they reach the market). Such a view belies the fact that, despite a massive trade surplus, the Japanese government is up to its neck in debt, and any government agency or ministry has to fight tooth and nail to win funds for new research by going through a long application process that begins in the summer and ends in the new year when the budget for the following fiscal year is announced (and even then the budget is subject to Diet approval).

MITI, the Science and Technology Agency and the Ministry of Education,

Last ditch attempt to save Messel pit

Munich

AN appeal to the United Nations has now been issued by West German palaeontologists in a last-ditch attempt to save the fossil-rich Messel pit. The pit, near Darmstadt, will become a rubbish dump this autumn under a ruling earlier this year of the South Hesse regional council, which has set aside 15 years of attempts by researchers to have the pit left alone.

The Messel pit, first excavated in the nineteenth century, is one of the richest known sources of well-preserved Eocene fossils in the world. Fifty million years ago, the area was a lake, in which dead animals and plants of the age were preserved in oxygen-poor sediments. The evolution of these sediments into oil shale led to the remarkable preservation of the fossils. Among the finds are a very rare, almost complete skeleton of a primordial cat-sized horse, *Propalaeot herium parvulum* and a large variety of bats.

Mining operations at Messel halted in 1971 when they became unprofitable. Talk of using the pit as a rubbish dump began soon afterwards. The government plans to dump up to 25 million m³ of rubbish into the 60-m-deep pit, thus solving a long-standing waste disposal problem in the Frankfurt-Darmstadt area.

Steven Dickman

Science and Culture have been quick to spot a potential money winner in the new superconductors. The agency set up a research committee in February with members drawn from the universities, government and industry to exchange information and hold symposia. Financial support for the committee's activities is provided by a small membership fee (¥200,000) from participant companies (which now number about 130). MITI also set up a similar "study group" which will publish a report on its meetings in September, and the Ministry of Education took the extraordinary step of extending a grant for research on oxide superconductors by advancing an extra ¥160 million (about \$1 million) (see *Nature*, 327, 356; 1987).

Now, with the season for budget applications little over a month away, MITI and the Science and Technology Agency have revealed their plans for fiscal year 1988. The agency proposes establishing a "multi-core project" to be carried out by four of its institutes: the National Research Institute for Metals, the National Institute for Research in Inorganic Materials, the Japan Atomic Energy Research Institute and the Institute for Physical and Chemical Research.

Each institute will form a research team in a particular field and these research 'cores' will interact by holding regular meetings, according to Masahiro Kawasaki, director of the agency's Research and Development Bureau.

Kawasaki says the agency will apply for ¥10,000–15,000 million (\$70–100 million) to set up facilities for superconductor research in the National Research Institute for Metals and National Institute for Research in Inorganic Materials in fiscal year 1988. Such a proposed level of funding is extraordinary as the total budget for these two institutes is normally only about ¥11,000 million, according to Kawasaki.

MITI's plans for government-funded research are more modest. The ministry intends to incorporate development of ceramic thin films and Josephson junction devices into an existing project. The "new functional devices" project aims at creating such things as three-dimensional integrated circuits and biochips and is supported with funds of about ¥1,400 million in this fiscal year.

MITI also hopes to take advantage of the present interest in superconductors to win ¥2,000–3,000 million in the next fiscal year to develop a 70,000-kW superconducting generator as part of the Moonlight Project. Moonlight was

created in 1978 during the last oil crisis to cut energy consumption through the development of a host of power-saving devices, including giant batteries for night storage of electricity, fuel cells and a Stirling cycle engine. Mitsubishi Electric Corporation and Fuji Electric Co. have already built a 30,000-kW superconducting generator under a different MITI programme which ended about three years ago, and Mitsubishi, Hitachi and Toshiba are expected to participate in development of the new generator which will employ niobium-titanium superconducting wire in the armature and a conventional magnet, according to Yasuo Hashimoto, manager of the planning department of Mitsubishi's Materials and Electronic Devices Laboratory. One hundred million yen (\$700,000) has been set aside in this year's budget for a feasibility study of the generator, development of which is expected to take about seven years.

But MITI's most ambitious plan is to establish a non-profit foundation, the "international superconductor industry and technology promotion centre", with funds from the private sector. MITI hopes 80 to 100 companies, including nine electric power companies and electronics, ceramic and cable manufacturers, will join the centre which could be set up as early as this autumn. The nine power companies already run a research cooperative, the Central Institute of the Electric Power Industry, which funnels funds to companies developing superconducting wire. And, according to Rizaburo Nezu of MITI's Agency of Industrial Science and Technology, this institute will play a role in establishing the new centre.

With support from 20–30 key companies, MITI hopes to set up a research institute within the centre which would hold international symposiums and exchange researchers with the US and Europe.

The Science and Technology Agency is also calling for international collaboration, and the agency's superconductor committee, which is open to foreign membership, is expected soon to be converted into a non-profit foundation and will hold its first international symposium towards the end of this year.

The Ministry of Education has yet to show its cards. Although the Ministry does not command the 'big science' budgets of the Science and Technology Agency and MITI, it does wield considerable control over research in the universities through the distribution of grants, and few researchers would risk accepting money from MITI or the agency for fear of being cut off from this source of funds. It is widely expected that Professor Shoji Tanaka, leader of the Tokyo University superconductor research group, will get about \$1 million when awards for this fiscal year are announced on 28 July.

David Swinbanks

Has Japanese technology made possible quiet Soviet subs?

Tokyo & Washington

Is there a direct link between the illegal export of milling machines to the Soviet Union by Toshiba Machine Co. and the appearance of quiet Soviet submarines? This question has been the subject of intense debate in the Japanese Diet during the past week, and it seems there is no simple answer.

Toshiba Machine's export of nine- and five-axis milling machine tools was a deliberate violation of COCOM rules that govern the transfer of Western technology to the Eastern bloc, and the general view in Japan is that the company should be punished. But there the consensus ends.

Foreign Minister Tadashi Kuranari declared in the Diet last week that there is a link between the illegal exports and quieter Soviet submarines. But other Diet members and military experts reject this view. They point out that quieter submarines belonging to the Victor III, Mike and Sierra classes first appeared in 1983 and early 1984 before completion of the Toshiba milling machines at the Leningrad shipyard in late 1984. The only quiet submarines to put to sea after 1984 were those of the Akura class which are built on the eastern seaboard of the Soviet Union far from Leningrad.

Hildeo Aoki, a defence commentator, further argued on the Asahi television network that the chief factor determining the quietness of Soviet submarine propellers is their design, not milling by Toshiba machines — a view that is shared by some US antisubmarine warfare experts.

Propellers produce noise as a result of cavitation. Streams of bubbles formed at the tip of the rotating propeller blade implode with a characteristic noise when they are drawn into the high-pressure region behind the propeller. Cavitation can be reduced by decreasing the speed of propeller rotation. The past several years have seen remarkable changes in the design of Soviet propellers to this end, and the latest models have huge multiple blades, shaped like scythes, that project far out of the water when the submarine surfaces.

Aoki also pointed out that cavitation accounts for only about 20–80 per cent of the noise made by submarines, other external sources of noise being the propeller shaft and the water-cooling system for the nuclear-powered engine (internal sources such as the turbine can be largely silenced through sound-proofing). If Soviet submarines have gone silent, other factors are involved apart from propeller design, he said.

Debate over the issue has caused con-

fusion in the Japanese government. Following Kuranari's assertion of a direct link (between machine exports and quieter Soviet submarines), Prime Minister Yasuhiro Nakasone said there is only "thick suspicion" of linkage, while International Trade and Industry minister Hajime Tamura said the link was "not yet clear". Just before Tamura's visit to the United States last week, the Nakasone administration reached a unified view of "thick suspicion", although chief cabinet secretary Masabaru Gotoda insisted that this unified view did not necessarily admit a link and "we will not admit a cause-and-effect relationship without conclusive evidence".

Upon his return from Washington, Tamura declared that he had been shown convincing evidence by the United States that Toshiba's machines helped silence Soviet submarines. But the information was classified. James Auer, special assistant for Japanese affairs at the US Defense Department, has made similar reference to top-secret evidence. But when pressed by Japanese reports to cite specific numerical data, he could say only

that Soviet submarines could previously be detected by underwater sonar at a distance of 80 km and this had been reduced to about 75 km. He admitted that there are other ways to reduce noise apart from propeller modification, but insisted that the Toshiba machines have been a key factor in silencing Soviet submarines.

On the other hand, former Navy secretary John F. Lehman Jr and Admiral James D. Watkins, recently retired chief of naval operations and a submariner, have blamed the spy ring headed by former Navy radioman John A. Walker Jr for the sudden Soviet advances.

Given the equivocal evidence that has been made public, why is the Nakasone administration so passive in its acceptance of US assertions that Toshiba's machines silenced Soviet submarines? Undoubtedly the Japanese government wants to calm the raging protectionist mood in the US Congress. But also the hullabaloo over Toshiba may be turned to advantage in the long run. It is no secret that Nakasone would like to strengthen US-Japanese military links, through, for example, participation in the US Strategic Defense Initiative (SDI). And Nakasone has already suggested increased collaboration in anti-submarine warfare activities following the Toshiba case.

David Swinbanks & Alun Anderson

Indian government cancels contract with US company

New Delhi

UNDER pressure from scientists, the Indian government last month took the unusual step of cancelling a multi-million dollar contract with the Hemlock Corporation of the United States for the manufacture of silicon. Under the contract signed in May 1984 with the Department of Electronics (DOE), the US company was to supply technology and equipment for setting up a \$100-million National Silicon Facility (NSF) to produce the 200 tonnes of polysilicon needed annually by the photovoltaic and electronics industry.

The termination of the contract has ended a three-year controversy over the choice of technology for NSF and has been hailed as a triumph of indigenous over imported technology. The step was also unusual because the Indian government had taken great pains to clinch the deal in the first place.

The Hemlock-DOE deal had been hanging fire ever since the contract was signed amid opposition from Indian scientists including Professor C.N.R. Rao, currently chairman of the Scientific Advisory Committee to Prime Minister Rajiv Gandhi. Critics had alleged that in awarding the contract to the company, the DOE had ignored the indigenous technology for

making silicon developed by scientists of the Indian Institute of Science (IIS) in Bangalore. In fact the IIS process was being used to make silicon in a 2.5-tonne capacity pilot plant at Mettur Chemicals, a private company. Despite an assurance by IIS and the company that it would be able to scale up the plant capacity at a fraction of the money charged by Hemlock, DOE had sought foreign collaboration with the approval of the then Prime Minister, Mrs Indira Gandhi. The reason, according to DOE, was that silicon produced by Mettur Chemicals was expensive and that its quality did not match that of silicon produced by the Hemlock process.

No other foreign collaboration had attracted so much criticism and the silicon deal with Hemlock became a test case in the continuing controversy over imported versus indigenous technology. Mr Rajiv Gandhi, soon after he became prime minister, was under pressure from some of his advisers to review the agreement with Hemlock. On his personal intervention, the DOE gave Mettur Chemicals a year to expand its capacity to 25 tonnes and to demonstrate the purity of the silicon it produced. The company fulfilled both conditions in March 1986.

K.S. Jayaraman

NASA explores its options for future space programme

Washington

A REPORT to be delivered to James Fletcher, administrator of the National Aeronautics and Space Administration (NASA), setting out potential new initiatives for NASA, will undoubtedly add fire to the debate over the direction of the US space programme. Prepared by astronaut Sally Ride, the report will present options ranging from a manned mission to Mars to renewed unmanned exploration of Earth. But the report emphasizes that there are certain steps NASA must take before any new programmes can reasonably be considered.

The Ride report attempts to answer the question of what it means to achieve leadership in space. Reaching the Moon first provided the United States with one type of space leadership, but unmanned exploration and advanced technology development has provided another, perhaps longer lasting, type of leadership.

In addition to the Mars and Earth options, the report examines a possible manned lunar base and unmanned planetary exploration. Along with the Mars and Earth-observing initiatives, these are intended to represent the spectrum of options for NASA. Although no option receives unqualified support, the lunar base and the Earth-observation options seem more reasonable short-term goals.

But the report stresses that virtually any activity NASA may undertake between now and the end of the century will require certain basic improvements. The launch fleet will have to be expanded beyond the shuttle, probably to include a

heavy-lift vehicle. There will also be a need for rockets designed to move payloads around once in orbit, and to carry them out of orbit.

New technology development is also critical. Apart from special application satellites, much of NASA's hardware is based on technology from the late 1960s and early 1970s. Space hardware for the 1990s must include more modern components. Renewed emphasis on life sciences in space is also critical, especially if longer-term missions are undertaken. The report also concludes that a space station of some description will be needed for both manned and unmanned missions.

The emphasis on Earth-observing missions will bring smiles to faces at the National Science Foundation, where Earth sciences have recently been reorganized under the banner of Global Geosciences. Those involved in the International Geosphere-Biosphere Program will see nothing amiss in NASA emphasizing Earth sciences. Planning has already begun at NASA for a fleet of four polar orbiting platforms, two provided by NASA, one by the European Space Agency and one by Japan. These platforms will be serviced by the space station, and used for Earth observations. There are also plans for five geostationary platforms.

The lunar base may be more controversial, but the Ride report sees it as a critical first step if Mars is to be an ultimate goal. Certainly it will be more expensive than enhanced Earth observation, although not as costly as a trip to Mars would be.

Joseph Palca

Enthusiasm for return to Mars

Washington

If the National Aeronautics and Space Administration is uncertain about whether returning to Mars in force should be a goal of the US space programme, such doubts did not exist for participants in this week's "Case for Mars III" conference in Boulder, Colorado. Nor was there doubt about continued Mars exploration during a unique, four-hour video conference between Soviet and US space scientists, engineers and 'explorers' sponsored by the Planetary Society. The video conference, dubbed Spacebridge, took place last Saturday, to coincide with the Mars conference.

About 25 participants from each country discussed issues ranging from the value of international cooperation in planetary exploration to engineering details for future missions. Soviet interest in searching for life on Mars surprised some US participants. Louis Friedman,

executive director of the Planetary Society, says part of the video conference was dominated by the life issue. He says the presumption is that by taking samples from further below the surface than those drawn by the US Viking landers warmer, wetter material might be found by Soviet probes. But Friedman adds that there is no clear evidence to suggest that such conditions would in fact exist.

There was no unanimity of opinion on the timing for a manned mission to Mars. Former US astronaut Buzz Aldrin argued that such a mission was achievable before the end of the century, but others felt that was far too optimistic. In mission design discussions, the possibility of cooperating on an engineering model of the martian surface was raised. Cooperation was also proposed as an efficient way of solving navigational problems related to landing on Mars.

Joseph Palca

British accident practice with radium

London

THE British Ministry of Defence admitted last week that it has been using α -emitting radioactive materials in nuclear weapons training exercises. The exercises, at undisclosed locations over several years, are designed to simulate accidents involving aircraft or ground vehicles.

Roger Freeman, Parliamentary Under Secretary for the Armed Forces, told a stormy House of Commons last week that liquids containing radium-233 (with a radioactive half-life of about 12 days) have been sprayed around an accident site to make the simulations "as realistic as possible". The ministry says that stringent safety procedures are followed, and that "there is no harmful effect".

The existence of the Nuclear Accident Response Organization, which has conducted the exercises, was revealed by *The Independent* newspaper. Some members of parliament complained at the use of any kind of radioactivity, but some researchers, accepting the need for "something to

You should count yourselves lucky we didn't use real missiles...



make the geiger counters tick", think it would have been better to avoid α -emitters. "A simple β - or γ -emitter would have been better", says Dr Murdoch Baxter of the Scottish University Research Centre.

The ministry says that radium-233 is used "typically" but not exclusively on training exercises, when it is mixed with barium sulphate and aluminium oxide. Contaminated soil is removed, and "after three or four months, no trace of radioactivity is left".

The hazards of the practice are disputed. Dr Barrie Lambert, lecturer in radiobiology at St Bartholomew's Hospital in London, says that the use of radium-233 is "unjustifiable", given the bone-seeking tendency of the material. But the ministry says that the maximum radioactive dose that could be received during an exercise is the equivalent of smoking a single cigarette (radioactive on account of polonium), which is estimated at 50 microsievert.

Kathy Johnston

Greens bested in West German uranium processing plant debate

Munich

THE conservative Environment Minister of the German *Land* of Hesse has won a publicity victory and embarrassed the Green Party by closing down part of the uranium-processing company Nukem. The minister, Karl-Heinz Weimar (Christian Democrat), who took over from Green Joschka Fischer in April, announced the partial closing on 3 July.

Nukem, in the city of Hanau, was criticized for inadequate fire protection and other safety deficiencies in the part of the plant converting uranium into fuel elements for use in research reactors. Weimar calls this "the heart of the operation" at Nukem, which carries out contracts for customers in Britain, France and the United States as well as West Germany. The facility will stay closed until Nukem has complied with the regulations — probably no more than three or four weeks.

But Weimar also announced that he has submitted a 36-point catalogue of potential problems, not made public, to Nukem and asked for a response by 3 August. This is thought to include even stricter safety requirements, which Nukem may be hard pressed to adopt at short notice. The outcome may be an even longer shutdown.

Nukem has nevertheless said it is willing to comply immediately with the fire safety requirement. A partial shutdown will not affect business, said a Nukem official. The plant can continue to operate until September using uranium already converted, but a shutdown for longer than a month would be "very difficult". All Weimar's directives so far concern the older part of the Nukem operation. A new section, begun in 1982, is expected to replace the older one in 1989 or 1990.

Weimar's quick and decisive action, which has won him much favourable publicity in the West German press, derives from his pragmatic attitude, and that of his party, to the nuclear industry. The ruling coalition, led by the Christian Democratic Union wants to use nuclear power at least until a cheap alternative comes along. Demanding that Nukem operate safely is one way to get it out of the limelight.

The Greens, on the other hand, have been trying for years to have Nukem shut completely. Even now, they complain that Weimar's measures are superficial. The Greens would at least like to see comprehensive protection against earthquakes, aircraft crashes and terrorist acts.

But the Greens did not push for safety improvements when they held the environment portfolio from December 1985 to April 1987, on the grounds that that

would have implied Green acceptance of the nuclear industry. The Green "fixation" on closing Nukem, said Green environment spokesman Chris Boppel, has led to the Greens' "self-destruction" in Hesse. The only Social Democratic-Green coalition in history broke up earlier this year over the licensing of the new section of Nukem.

The Greens nevertheless take solace that they "started the avalanche" of public concern about Nukem in the early 1980s. Their complaint that Nukem and its sister company Alkem had been operating illegally since 1975 has led to a lawsuit brought by the Hanau district attorney's

office against the president of Alkem, another Alkem manager, and several former Social Democrat (SPD) officials. The defendants are charged with making six changes in Alkem procedures outside the guidelines of a "preliminary licence" granted by the SPD government in 1984. The trial begins on 10 August.

Another pending lawsuit may have grave consequences for Nukem. According to Green spokesman Boppel, the Hanau district attorney is planning to indict three former cabinet ministers on charges of colluding with Nukem managers.

The district attorney's office has refused to reveal the defendants in the case until the indictment becomes final. Whether it can be proved that Nukem and Alkem have operated illegally is open to question.

Steven Dickman

Canada aims to join remote sensing field with new satellite

Toronto

CANADA plans to go ahead with its first ambitious remote-sensing project — a radar sensor due to be put into an 800-km Earth orbit in 1994. So much is clear from the announcement by science minister Frank Oberle that many financial details of the C\$725 million Radarsat project have now been settled. Formal cost sharing agreements with the United States and Britain should be completed by late fall.

It was no accident that Canada was among the first countries in the world to embrace the space age. As a large country dependent on natural resources, with a small, dispersed population and many remote, undeveloped regions, it was tailor-made for the benefits of space-based technologies such as communications and remote-sensing satellites. In 1972, the same year Canada became the first country to put a communications satellite in orbit exclusively for domestic use, it built the first non-US ground station to receive Landsat data.

In the following 15 years, Canadian companies became leaders in ground-based receiving and image-processing and data-analysis technologies, capturing an export market worth some C\$60 million annually through sales and services to more than 100 countries.

Radarsat will be able to collect data at night and through clouds — critical for reliable coverage of Canadian territory, especially in the north — by using an advanced synthetic aperture radar, a technique that uses the spacecraft's own motion in processing radar signals reflected from Earth. Radarsat will have multiple, adjustable beams able to resolve objects from 100 m to 10 m in size, smaller than the US Landsat and the French

SPOT. The satellite will be used to collect data on crops, forests, mineral resources, Arctic ice, sea state and weather conditions. The Canadian government also views it as an important tool for monitoring shipping in Arctic waters and asserting sovereignty in the north.

Canada is taking the lead in the 10-year, C\$725 million programme, but Britain and France are expected to provide some 40 per cent of that total in return for a share of the data, according to Oberle. British Aerospace plc will provide the spacecraft structure, worth an estimated C\$197 million, and the United States will provide launch services. Canada will build the radar and the data-collection system and will manage data-gathering and analysis.

Oberle says the private sector is planning to establish a consortium to market and distribute the data, which will be sold to Canadian federal and provincial government agencies, foreign countries, corporations and individuals. The sale of data is expected to defray part of the C\$10 million annual operating costs. If projected sales of C\$150 million over the satellite's five-year lifetime materialize, the consortium will return C\$34 million in royalties to the government.

When the Canadian cabinet first considered its budget, the project was competing with two other big items — M-Sat, a mobile communications satellite, and the robotic servicing system Canada plans to contribute to the US/international space station. Decisions in favour of the other two projects all but orphaned Radarsat, and its survival depended on the ability of proponents to put together a new coalition of public, private and international support.

Lydia Dotto

Barry Jones's science ministry scrapped as Hawke reorganizes

Sydney

AUSTRALIA's Department of Science has been scrapped. In a major reorganization of Australia's public service with a reduction of the number of ministries from 27 to 16, science has been divided between new superportfolios. The move was the first action of Prime Minister Bob Hawke after the re-election of his Labor government on 11 July.

Apart from gains in efficiency, which Hawke estimates at A\$96 million a year, the reorganization also helps to solve his problem of pressure from the back bench for promotions. Despite the reduction in the number of ministries, the number of ministers will increase from 27 to 30; 16 of them will have cabinet rank, each heading one of the mega-departments, and the remainder will be junior ministers assisting them.

Before the prime minister's announcement, the names of both Australia's flamboyant Minister for Science, Mr Barry Jones, and the Minister for Education, Senator Susan Ryan, were on a list of four ministers said to be in danger of relegation. They owe the weakness of their position to the over-representation of their centre-left faction in the old ministry.

Jones was among the audience at the prime minister's press conference. What he heard was that much of the Department of Science he has headed is to be taken over by the Department of Industry, Technology and Commerce (DITAC), which will be the new home for science policy, the Commonwealth Scientific and Industrial Research Organization (CSIRO) and the Australian Nuclear Science and Technology Organization, as well as the patents and national standard divisions.

Research finance, almost certainly in the form of the Australian Research Council recommended by the recent review of Australia's research policy (see *Nature* 326, 431; 1987) and whose establishment was endorsed by both sides during the election campaign, is bound for a new Department of Education, Training and Employment.

Jones has said that he has no idea which job he will end up with in the new ministry. It seems certain that he will not be put in charge of DITAC, the province of Senator John Button, while Hawke has already made it clear that none of the major economic portfolios will change hands. Although Button is also a member of the 'centre-left faction', he is the leader of the government in the Senate and probably the number three man after the prime minister and the treasurer, Paul Keating.

The reorganization makes complete his

takeover of the science and technology portfolio, given to Jones when Labor came to power in 1983. The technology component was annexed by Button after the previous election in 1984. Mr John Dawkins, at present Minister for Trade, is tipped for the Ministry of Education, Employment and Training.

Jones has many supporters. Both the CSIRO Staff Association and the Federation of Australian Scientific and Technological Societies have sent telegrams to the

prime minister in praise of his unprecedented energy and enthusiasm as Minister for Science, as well as for his work in elevating science to the mainstream in Australia and for his advocacy of a brain-based economic recovery for the country.

Lumping science together with industry, technology and commerce is indicative of the government's continuing intention to make science pay its way. The new public service structure could work in favour of science. There will be representation in the cabinet from two separate ministers, something science never had in the previous government, in which Barry Jones was not a member of the cabinet.

Charles Morgan

France tries to create and promote new technology

Paris

THE French government is increasingly alarmed that France is falling behind its European partners in its ability to create new technology. With 12,000 patent applications a year in France, compared with 20,000 in the United Kingdom, 30,000 in West Germany and 60,000 in the United States, the need to turn fundamental research findings into marketable products is seen by the government as acute.

To help to redress this imbalance, the prime minister, Jacques Chirac, has announced special measures to encourage private-sector research, ranging from tax incentives to financial rewards for researchers who leave state-financed research for industry.

The problem, according to the government, is that the large state-supported research bodies (the *grands organismes*) not only cream off the best scientists but have fostered a closed system which has little contact with industry, either to define areas of need or to develop and commercialize spin-off from research.

To facilitate the transfer of knowledge from research centres to industry, an interministerial 'new materials' programme is now to be set up, linking defence, industry, telecommunications, the atomic energy commission (CEA) and the *grands organismes*, such as CNRS (the Centre Nationale de la Recherche Scientifique). The programme will operate with a FF200 million budget, provided from the respective partners' existing funds. Also, to 'irrigate' the private sector, researchers will be offered the equivalent of a year's salary to leave state-supported institutes and join industry.

The Agence Nationale pour la Valorisation de la Recherche (ANVAR), a quasi-autonomous non-government organization set up in 1979 to provide interest-free loans for innovative research and to encourage links between industry and

fundamental research, is to receive a 10 per cent increase to FF917 million in its budget for 1987-88. This is a dizzying *volte face* for the government, which cut ANVAR's budget by 40 per cent on coming to office last year.

Some critics are doubtful whether the 'new materials' programme will be an effective focus for collaboration between science and industry, while the 'bribing' of researchers to join industry was tried by the Socialists, only to be dropped when inadequate research resources in the private sector attracted few takers, a situation unlikely to change radically even with further tax allowances for research spending. Meanwhile, scientists in the *grands organismes* are waiting to see if the handouts to industry will be paid for by cuts in their own budgets when parliament reassembles in the autumn. Peter Coles

AIDS tests in India

New Delhi

All foreigners, except journalists and diplomatic staffs, intending to stay for more than a year in India will now have to be tested for AIDS (acquired immune deficiency syndrome). The move, announced by the health ministry, is said by officials to be an attempt to placate foreign students, mostly from Africa, who complain that India's policy of testing students (now relaxed to apply only to newcomers) is discriminatory. K.S.J.

History of science

London

A working party of the British University Grants Committee has produced an encouraging report on the future of the history of science in British universities, saying that it has a "service function" that has become "more, not less, important". It recommends continuation of the existing

Vaccines developed in United States to be tested in India

New Delhi

INDIA will become the testing ground for several new vaccines being developed in the United States, following an Indo-US agreement last week on a Vaccine Action Programme (VAP) "to develop and test vaccines and diagnostic techniques for major communicable diseases".

The agreement, signed by Mr John Gunther Dean, the American Ambassador, and Dr S. Ramachandran, Secretary of the Department of Biotechnology, has paved the way for trying out in India advanced and genetically engineered vaccines that might, for practical reasons, be difficult to test in the United States.

India has always been sensitive on the issue of its people being used as guinea-pigs for the trial of drugs and vaccines developed elsewhere, which is one reason for the delay of nearly 18 months in signing the agreement. This difficulty has now been met by the United States, which has associated Indian scientists with the US laboratories where the vaccines are being developed.

What form this association will take is not clear; most of the US vaccines are already in an advanced stage of development and waiting only for clinical evaluation. But by projecting the vaccines as products of Indo-US collaboration, the United States hopes to avoid opposition to clinical trials in India.

Among the vaccines to be tested are those against diarrhoeal diseases (rotavirus, cholera, shigella, *E. coli* and salmonella), a cellular pertussis vaccine, an oral typhoid vaccine and a recombinant

departments at Cambridge, Edinburgh, Manchester and Sussex, expansion (if anything) at Bath and Lancaster, restructuring at Manchester, Oxford and London ("where two or three potentially major centres have been allowed to run down to the point at which they are barely viable") and the continuation of most other activities in the field. □

Ocean resources

London

Resources from the oceans could provide a world market worth more than £160,000 million a year, according to series of studies commissioned by the UK Department of Trade and Industry. The documents are mostly concerned with deep-sea mineral reserves, offshore petroleum and seabed mapping. While British continental shelf resources are not estimated to be large, the reports (which have cost £300,000) say that there may be a demand for British skills. K.J.

DNA vaccine against hepatitis-B. The collaboration also provides for the testing of the vaccinia/rabies glycoprotein recombinant vaccine developed at the Wistar Institute and used in a controversial experiment on cattle in Argentina (see *Nature* 324, 610; 1986).

The inclusion of the Wistar vaccine in Indian trials has raised eyebrows among a section of the medical community, but Ramachandran says the Argentine experience will not be repeated in India. He says that the Argentine controversy arose because Wistar failed to inform the government. He said that there are 7 million dog bites a year in India; it is proposed to control rabies in stray dogs by mixing the recombinant vaccine with dog food. In Argentina, the vaccine was inoculated.

The five-year project is estimated to cost \$9.6 million, out of which the US government will contribute \$7.6 million. Besides the Department of Biotech-

nology, the Ministry of Health and the Indian Council of Medical Research (ICMR) will implement the project.

Allaying fears that the agreement with the United States will make Indians guinea-pigs, Ramachandran said the rights and welfare of human subjects of research will be protected, taking into account the laws and regulations in both countries. Clearance would be obtained from the drug controller of India before tests of the vaccines, he said.

Whether or not the United States will succeed in testing its vaccines depends on the ICMR, which advises the drug controller whether a vaccine should be tried on humans. Dr A.S. Paintal, director general of ICMR and a stickler for medical ethics, said he would not allow any vaccine to be used in India unless it was also approved for use in the United States by the Food and Drug Administration.

As for the US motive in collaborating with India, Paintal says it represents only the natural curiosity of scientists to find out the efficacy of their vaccines by testing them in places where the diseases exist.

K.S.Jayaraman

British compromise with Europe over research programme

London

SCIENTISTS involved in projects supported by the European Community's research programme are relieved that the long-running feud between Britain and its partners has been virtually resolved. But, after months of blocking by Britain, the compromise programme has a far slimmer budget than originally proposed.

Community ambassadors have agreed on a five-year programme costing 5,200 million ECU (European Currency Units, 1 ECU = £0.69), 417 million ECU less than the amount supported by all other 11 member states in March. Britain is still formally blocking the extra 417 million ECU, but will agree before the end of the year on condition that progress is made on curbing other expenditure, especially on agriculture. An additional 800 million ECU will be available for commitment during the next five years, but cannot be spent until after 1991.

The compromise came after Britain changed its definition of "existing expenditure" on the programme. The Prime Minister, Mrs Margaret Thatcher, had told last month's summit that Britain would agree to maintain the research at the level of existing expenditure so that important projects could continue; originally, the British government had argued that existing work was equivalent to 4,000 million ECU only.

The programme will be drawn up as though the full 5,600 million ECU is

available, which observers see as an indication of Britain's willingness to compromise.

Ministers are expected to give formal approval to the programme this week. The European Parliament will discuss it in September, and no funds will be officially released until after that. The parliament had threatened to withhold funds in the areas of telecommunications, medical research and science and technology for the developing world unless the overall five-year framework was agreed.

The compromise should mean cooperative research programmes such as Esprit (information technology) and Race (broadband telecommunications), as well as others, can continue without further interruption. But some research has already suffered because of the long delay — the European Commission says the second phase of Esprit could be set back, with some researchers temporarily laid off, and the European Parliament Research Committee has been told that British Telecom has had to disband its telecommunications research team.

Some scientists, puzzled by Britain's long battle against the research programme, point out that for every £1 Britain puts into joint community research, up to £1.30 comes back in contracts.

"After all this delay, everybody in Europe has lost", says European member of parliament Glyn Ford, a member of the research committee. Kathy Johnston

White rule in South Africa

SIR—It would take too long to take up every piece of explicit and implicit racism in the article on the boycott of South Africa (*Nature* 327, 259; 1987). Instead I would like to outline some arguments for the scientific boycott of South Africa.

First, far from being a "scandal", there is a perfectly good moral case for excluding South Africans; those whites who choose to enjoy the advantages of living comfortably in South Africa, at the expense of the black population, are party to denying the vast majority of the South African population "the right to contribute to research", purely on grounds of race. It is, therefore, only right that they should be denied the benefits of attendance at international conferences and so on. No doubt some white South African scientists are secretly working for the overthrow of the present system, but no doubt they would equally secretly welcome the boycott that seeks to hasten that overthrow.

Apartheid is conveniently taken to be the whole issue, but it is not. It is only one form of white domination, which did not begin all of a sudden in 1950. John Maddox's "reformers" are those who think it might no longer be the best way of keeping the mass of the black population in subjugation. White rule in South Africa is a system of colonial settler domination, established and maintained by force. Typically, Maddox refers to "bombings" by the African National Congress, but is silent about the South African state's continual acts of aggression against its neighbours and against the mass of its own people. White South Africa is making war on the black population. That is what "reformer" Professor Cloete's right of South Africa to defend itself means. In response, the people of South Africa are beginning their war of liberation; which brings me to the more important case for a scientific boycott — to weaken the power of the white South African state to wage war against the people.

Most readers of *Nature* would agree that scientists are rather valuable to a country, more so perhaps than singers or rugby players. It is unnecessary to argue the case that science is essential to the production of weapons, to the supporting industries and to the economy that sustains a war effort. The article itself makes it abundantly clear that science in South Africa is utterly dependent on the international scientific community. Maddox is effusive in his praises of (white) South African science, but if South Africa were totally cut off, world science would hardly notice, while South African science would be, as one of Maddox's informants puts it, "dead". This process of isolation would also weaken the regime by encouraging native white South African scientists (and

other intelligent people) to leave the country for the duration.

I presume Maddox is one of those who favour 'peaceful change' in South Africa, by which they mean that the white rulers should come up with some less offensive system without being subject to any serious coercion. In fact, the only chance of a more-or-less peaceful solution is for the white supremacists to realize that they must lose, and accordingly negotiate a surrender of power. That is why all humane scientists should support measures such as the scientific boycott which contribute to so weakening South Africa that even those dinosaurs recognize their day is done.

The profound racism of Maddox's approach is evident from his outrage at any restriction on the right of 6 million white South Africans to take part in scientific research, together with his complete indifference to the fact (evident even from his article) that 17 million black South Africans are denied the possibility of a scientific education and career. We are often told that the issue of South Africa raises complicated and difficult questions. In fact, in essence there is only one question — which side are you on? The side of the racists, seeking to maintain their domination by military force, or the side of the people, seeking to establish a non-racial democracy? I very much hope that on this issue *Nature* is not the voice of British science.

J.G. WILSON

8 Agincourt Road,
London NW3 2PD, UK

University sports

SIR—I hope that British universities will learn from the experience of their transatlantic counterparts (*Nature* 325, 470; 1987) the lesson that they are likely to get badly burned if they attempt to introduce professional or quasiprofessional athletics on the US model.

There seem to me to be only two sensible ways of handling intercollegiate athletics, and, in general, US universities adopt an unsatisfactory mixture of the two that gives them the worst of both worlds.

One method is to call things by their right names and employ a professional football or basketball team, whose members are paid high salaries but are in no sense students at the university. In a few years, they should obviously be able to earn enough to put themselves through college if they could meet regular academic requirements.

The other method is to admit all students for whom there are places and who meet the academic requirements of the university, and then choose teams

from them. This is approximately the system that I assume still exists in British universities.

The unfortunate aspect of the present system of recruitment for intercollegiate athletics in the United States is the double educational standard that it implies. For example, if students in general require a Scholastic Aptitude Test (SAT) score of, say, 1,100 for entry to a university, athletes will probably be admitted with SAT scores of 700 to 800. In addition, the temptation for enthusiastic alumni to 'sweeten the pot' (illegally, of course) by paying for cars and apartments for athletes in order to induce them to come to the *alma mater* appear to be quite irresistible.

This leads to the kind of unpleasantness that has recently been seen at Southern Methodist University in Texas, which has effectively been banned from intercollegiate athletics for one year for illegal payments to players or the earlier stupidity at the University of Georgia where a faculty member was fired for failing athletes academically, and had to sue the university to get reinstated. She won, and the president of the university resigned.

In short, I urge you to use your influence to prevent British universities from starting down the slippery slope of giving some students easier academic options in the hope of getting a better football team or a faster boat crew. Universities in the United Kingdom have been treated in a stupid and niggardly fashion by the government but there is no need for them to compound their problems by getting drawn into the mess of semiprofessional athletics. It will raise almost no money for educational purposes, and they will get involved in a nasty mixture of half-truths and evasions about who is or is not a student eligible to play for the university.

Let them keep out of it altogether and concentrate their energies and resources on higher education.

CECIL S. CUMMINS

Department of Anaerobic Microbiology,
Virginia Polytechnic Institute
and State University,
Blacksburg, Virginia 24061, USA

Bad example

SIR—What hope have we of getting the facts through to the general public when *Nature* gets it wrong ("Keeping politics out of AIDS" *Nature* 327, 447; 1987)? You talk about the spread of AIDS and infection with AIDS; it is the virus (HIV) which is infectious, not the syndrome. Maybe I am being pedantic, but despite extensive education campaigns there is still considerable confusion over the facts about AIDS. Please don't add to it.

RUTH H. WALKER

Scottish AIDS Monitor,
PO Box 169,
Edinburgh, UK

Can the Universe be closed?

Astronomers hanker after the notion that the density of the Universe is the minimum required to prevent indefinite expansion, but the evidence is not all on their side.

ONE recurring theme in cosmology is that the density of the Universe should be exactly equal to the critical density separating universes that expand forever from those that eventually collapse back on themselves. Oddly, there has never been any astronomical evidence that this is in fact the case, and indeed there is plenty of evidence that the density is less than the critical value. But there is enough leeway in the observations to accommodate the prejudice that the value of Ω , the ratio of the cosmic density to the critical value, really ought to be one.

The traditional way to find Ω is to measure the mass of as many different things as possible (galaxies, groups of galaxies, clusters of galaxies), count how many of them can be seen and then to do the appropriate multiplications. The simplest route to the mass of a galaxy is to make some assumptions about the kinds of stars it contains, when its light output is a measure of its mass. But the masses of galaxies can also be derived from their dynamical interactions, and the clear result is that most objects in the Universe are heavier than they look. Several arguments suggest that galaxies carry with them haloes of unseen dark matter.

A widely accepted, but more arguable, extension of this view is that the ratio of dark to visible matter (familiar as the mass-to-light ratio M/L) is greater for larger groups of galaxies, implying that the best estimates of Ω will come from observations of the richest clusters. The lore at present is that the mass-to-light ratio on these largest scales is one or two hundred times as great as the corresponding value for the Sun. If Ω is to be one, this ratio has to be several hundred, perhaps even a thousand, but on the other hand, it is by no means clear that M/L as a function of scale has reached an asymptote even in the richest clusters of galaxies. There is still room for $\Omega = 1$.

But now there may be a little less room, according to L. Cowie, M. Henrikson and R. Mushotzky (*Ap. J.* 317, 593; 1987). A direct way to determine the mass of a rich cluster is to measure the density and temperature profile of the hot gas that is smoothly distributed between the galaxies and which is recognizable by the X-rays it emits. The gas traces the gravitational potential, which is related to the mass distribution by a Poisson potential equation.

The difficulty is that the temperature profile of the gas cannot be deduced

directly; many previous estimates of cluster mass have been based on the assumption that the gas is isothermal, which does not fit the X-ray observations. Cowie and his colleagues do not claim to have an exact model of the gas distribution, but by taking various models, and checking that they can fit the optical data, they find that the mass-to-light ratios of clusters need be only a little greater than 100 and may be as small as 60. It is no surprise that better modelling should lead to altered values, and the change is not huge. Its significance is that the density needed to close the Universe may be five or ten times the density of dark matter even in the richest clusters.

Even if absolute determinations of the density of clusters were possible, they would not unambiguously reveal the overall density of the Universe, which is why cosmologists have been looking for ways of finding Ω directly. The increasing refinement with which the Universe is mapped has allowed some intriguing if not yet convincing estimates. Thus, if gas can trace the gravitational potential in a cluster, why should not galaxies trace the gravitational potential in the Universe?

Improved astronomical techniques have delivered the redshifts as well as the positions of large samples of galaxies. Redshift translates into distance according to Hubble's law, so that the catalogues of galaxies listed by their two-dimensional position on the sky have been turned into catalogues of galaxies in three-dimensional space. So why not divide the mass by the volume to obtain large-scale Ω ?

There is a circular argument to unravel at the outset. If the galaxies were distributed regularly, the Universe would expand regularly and redshifts could be converted precisely into distances. But in reality, galaxies are clustered, and their velocities are affected by the dynamics of clustering on top of the cosmological expansion, so that their redshifts are contaminated by their peculiar motions, which are in turn of crucial importance in tracking the large-scale dynamics of the clusters... This seems an impossible tangle: to get the positions of galaxies, the peculiar velocities must be subtracted; but to find the peculiar velocities, the spatial distribution of the galaxies must be known.

Luckily, a little ingenuity allows some unravelling. In a dense cluster, galaxies have a characteristic velocity distribution that is independent of position; each

redshift contains a random component so that, when redshift is taken as a proxy for distance, the cluster appears to be elongated along the line of sight. But, for widely separated galaxies, it has usually been assumed that the peculiar velocities are of no significance because the Hubble velocities are so much bigger. N. Kaiser has now shown (*Mon. Not. Roy. Astr. Soc.* 227, 1; 1987) this plausible reasoning to be false, at least in one important case.

Kaiser defines the acceleration vector as an integral over a specified volume weighted according to the density fluctuations of the galaxy distribution. Then, at a point within the volume, the peculiar velocity induced by gravitational non-uniformity is the product of the acceleration vector and a known function of Ω . So, by measuring the galaxy distribution in our vicinity and knowing (as we do) our velocity relative to the microwave background, Ω should be calculable.

Sadly, Kaiser shows that this procedure does not work. Estimating the local acceleration vector requires that the positions of all neighbouring galaxies should be known, and it has been customary to take redshift as a measure of distance. But when Kaiser corrects the integral for the acceleration vector for the peculiar motion of local galaxies, the correction is disconcertingly comparable in size with the vector itself — unless Ω is much less than one.

This result is disappointing but not surprising. Both the acceleration vector and the peculiar velocities arise from departures of the large-scale galaxy distribution from uniformity. Kaiser lessens the disappointment by describing a more reliable, if more difficult, way of finding Ω on a large scale.

The proposal is that the statistics of galaxy distribution should be measured separately along and across the line of sight. Provided that the underlying distribution has no preferred direction, the difference between the two should represent the contamination by peculiar motion. A sufficiently precise analysis should yield a value of Ω directly related to large-scale cosmological dynamics. As if to sustain the spirits of the " Ω must be one" school, Kaiser also says that a simple modelling of peculiar velocities in the (local) Virgo cluster is not inconsistent with a Universe in which the critical density is exactly one. It is equally not inconsistent with $\Omega = 0.2$. **David Lindley**

Radioastronomy

Refraction in the Galaxy

Antony Hewish

THE interstellar medium, a complex mixture of hot, low-density plasma and cool, dense clouds of neutral gas, is repeatedly stirred and injected with fresh material by supernova explosions. The ionized component fills the greatest volume and information on its fine-scale structure, based on observations of pulsar scintillation, has been steadily accumulating over the past two decades. An unexpected result is the recent observation¹ that localized clouds of much higher density can occasionally produce enough refraction to occult, or obscure, extragalactic radio sources detected at frequencies as high as several gigahertz. Although the detailed physical nature of these regions is unresolved, it is clear that radioastronomers may sometimes encounter considerably worse 'seeing' through the Galaxy than they previously expected. On page 324 of this issue, R.W. Romani *et al.*² propose one possible explanation of these effects.

The line-of-sight to a distant radio source traverses various plasma regions including the ionosphere, the solar wind, the interstellar gas and the intragalactic medium, all of which except the last are known to contain inhomogeneities that phase-modulate incoming wavefronts. The first use of propagation effects to study astrophysical plasmas was in 1952 when an attempt was made to measure the density of the solar corona far from the Sun. The corona should act as a giant diverging lens, producing a bright caustic ring surrounding a zone of zero intensity. The Crab Nebula passes within a few degrees of the Sun each year during June and an occultation was predicted. In the event, what was observed was a 'bathroom-window' effect in contrast to lens-type action; the random refraction by small-scale irregularities produced multiple images and a broadening of the angular size of the source without loss of intensity. Later it was realized that the illumination of these plasma irregularities by a very compact radio source could be sufficiently coherent to produce speckle patterns at the Earth. This phenomenon (interplanetary scintillation) is useful for measuring the speed of the solar wind, mapping interplanetary transients and estimating the angular sizes of compact radio sources at low frequency (~ 0.1 GHz).

Following the discovery of pulsars, analogous scintillation effects were soon recognized, but on a much longer time-scale. These were caused by plasma clouds, about 10^9 m in diameter, in the interstellar medium. Because interstellar winds are relatively slow, the drift speed

of these speckle patterns is dominated by the motion of the source and observations showed that pulsars move at high speeds, which is consistent with their escape from close binary systems. In one notable case scintillation was used to estimate the orbital speed of a pulsar that is still gravitationally bound to its binary companion³.

Interstellar scintillation has been used to probe the hierarchy of plasma clouds in the Galaxy, showing that the wave-number power spectrum follows a Kolmogoroff law (also found in atmospheric turbulence) on scales near 10^9 m. It was suggested⁴ that this spectrum extends to at least 10^{14} m; but this turns out to be an oversimplification. When pulsar intensity variations were studied as a function of their radiofrequency and time, remarkable quasiperiodic patterns became evident. The random speckle patterns caused by small clouds were being superimposed by refraction through larger clouds to produce interference fringes⁵. This means that the densities in the larger structures exceed the Kolmogoroff law, and the occultations measured by Fiedler *et al.*¹ demonstrate the same conclusion even more dramatically.

The combined effects of diffraction and refraction on waves propagating through a hierarchy of cloud sizes are not easy to analyse theoretically. Recent work has invoked catastrophe theory in the treatment of irregular caustics such as occur in the

bright tracery patterns seen on the bottom of a swimming pool in sunlight. An alternative explanation of the occultation of quasar 0954+658 in terms of caustics is now advanced by Romani *et al.*² in this issue. In contrast to Fiedler *et al.*, who propose a model involving multiple imaging from a random population of clouds in some localized region, Romani *et al.* suggest that the observations can also be explained by caustics formed by refraction through a single, larger cloud. The latter requires a higher column density along the line-of-sight, and suitable structures might be found within old supernova remnants. If the clouds were sheet-like, Romani *et al.* suggest that occasional dramatic lensing effects would occur when the sheets are viewed edge-on, whereas they would otherwise produce the usual scintillation effects.

Whether the occulting regions act like bathroom windows or diverging lenses, observers have been alerted to a new and severe type of bad seeing through the Galaxy that can restrict for example, pulsar timing measurements or the mapping of extragalactic sources at high angular resolution. On the other hand, the existence of these dense plasma clouds provides new information on the physical state of the interstellar medium. □

1. Fiedler, R.L., Dennison, B., Johnston, K.J. & Hewish, A. *Nature* **326**, 675–678 (1987).
2. Romani, R.W., Blandford, R.D. & Cordes, J.M. *Nature* **328**, 324–326 (1987).
3. Lyne, A.G. *Nature* **310**, 300–302 (1984).
4. Armstrong, J.W., Cordes, J.M. & Rickett, B.J. *Nature* **291**, 501–504 (1981).
5. Hewish, A., Woźniak, A. & Graham, D. *Mon. Not. R. astr. Soc.* **213**, 167–179 (1985).

Antony Hewish is Professor of Radioastronomy at the Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, UK.

Animal behaviour

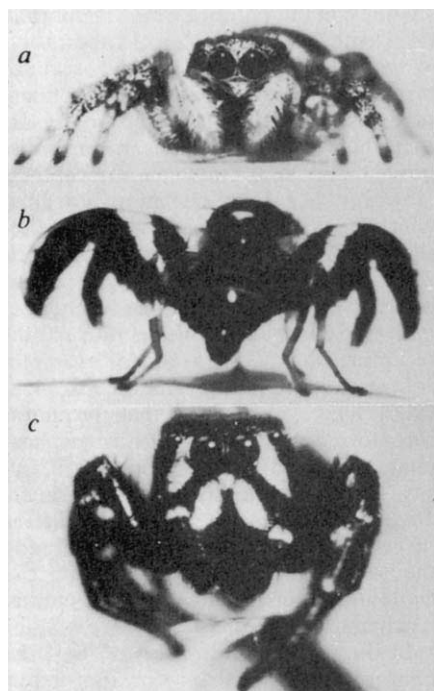
The fly that traps the spider

M.F. Land

JUMPING spiders identify other members of the family by the characteristic pattern of their legs. Oskar Drees showed in 1952 that although jumping spiders will pounce on a wide variety of small moving objects, models with a pattern of oblique lines on either side of a 'body' are not attacked, and may indeed evoke sexual behaviour from males (*Z. Tierpsychol.* **9**, 169–207; 1952). If this method of identification is as simple as it seems, then one way for a potential prey to avoid capture would be to provide itself with a pattern of lines that resemble the legs of jumping spiders (see figure). It seems that there are some flies that do indeed protect themselves in this way. Two recent papers in *Science*, one by M. H. Mather and B. D. Roitberg and the other by E. Greene *et al.*, describe how tephritid fruitflies prevent jumping spider

attack by a combination of patterns on the wings, and a zig-zag dance they make when disturbed, which resembles the brief agonistic display that jumping spiders often make when they meet.

Mather and Roitberg (*Science* **236**, 308–310; 1987) used the snowberry fly *Rhagoletis zephyria*, whose wings are strikingly banded and convincingly spider-like when held in the proper posture. They find that jumping spiders (*Salticus scenicus*) are much less likely to attack these flies than either houseflies or tephritids whose white wing bands had been inked over. The spider often actually flees from the fly, as it would in a territorial encounter with a conspecific. Mather and Roitberg show that although the wing-pattern is essential, it is really effective only when the wings are elevated and



a, Front view of zebra spider *S. scenicus* and b, posterior view of snowberry fly *R. zephyria*. (Courtesy of M. H. Mather and B. D. Roitberg.) c, Male jumping spider *M. imperialis*.

the fly performs its side-to-side dance.

The second paper, by E. Greene, L.J. Orsak and D.W. Whitman (*Science* 236, 310-312; 1987), reaches essentially the same conclusions about a New Mexico tephritid *Zonosemata vittigera*, but these authors took the study rather further. They argue that the real reason for the resemblance between the fly and spider might be because jumping spiders are themselves unpleasant, and that mimicking them would protect the flies against other predators that would normally avoid jumping spiders (Batesian mimicry). They tested tephritids and houseflies with four other predators: an oxyopid spider, an assassin bug, a praying mantis and a whiptail lizard, and found that there is no difference in the acceptability of the two flies, measured as the time taken to capture them. This result contrasts with the situation with jumping spiders, which back away from the tephritids but attack houseflies instantly. Thus, the protection afforded by the tephritid wing pattern and display does seem to be specific to predation by jumping spiders, and other predators are not fooled.

A problem in trying to work out the origins of this evidently successful strategy is that the display that tephritids make to jumping spiders is the same as that used in courtship, according to Mather and Roitberg. Many flies, especially those with patterned wings, have wing-waving displays whose function is to attract the attention of conspecifics, and males often display to anything that moves. It could be, as Green *et al.* suggest, that the tephritid mimicry evolved from courtship

Infrared excess stirs cosmologists

ELECTROMAGNETIC radiation of all wavelengths fills the Universe. The 3-K microwave background is the best known and most studied cosmological emission, but in the past decade or so, sophisticated (often airborne) instruments have enabled astronomers to measure the temperature of the heavens at almost every wavelength. Much of this radiation is simply accumulated emission from our own Galaxy, but there are a few windows in the spectrum through which the Universe can be glimpsed. At a recent meeting*, signs of an excess of infrared at two different wavelengths were reported, and theoretical cosmologists were quick to hang their own theories on the proffered pegs.

T. Matsumoto (Nagoya University) described results from an infrared detector put above the interfering atmosphere by rocket. Such experiments are both technically difficult, and hard to interpret theoretically. The detector must be cooled, usually by liquid helium, so that thermal emission from the instrument itself does not overwhelm the tiny cosmological signal. And the detected signal itself will usually be dominated by radiation from various local sources: integrated starlight, zodiacal light, emission from interplanetary dust, and so on. All this must be independently measured and subtracted before any putative cosmological radiation is revealed.

Matsumoto first drew attention to measurements in the near infrared (wavelengths of a few micrometres), and claimed that there was evidence for a small excess at about 2 μm . On this point the audience seemed generally unconvinced; most participants were inclined to think that a small adjustment in the various subtractions could easily account for the supposed signal. But Matsumoto also showed results from a more recent flight, in which there

was credible evidence for an excess in the far infrared, at a wavelength of a little less than 1 mm. For this wavelength, at a higher energy than the peak of the 3-K background and less than the characteristic 20-K emission from galactic molecular clouds, the subtraction is easier, and the proposed cosmological origin of the excess seems the most likely explanation.

Where does the signal come from? P. L. Richards (Berkeley), talking about the microwave background, said unequivocally that there was now no evidence for any distortion from an exactly thermal spectrum at 2.7 K. But Matsumoto's signal would correspond to a component of thermal emission at 3.1 K. Other explanations all attribute the excess to some kind of comparatively recent astrophysical process. In fact, as B. J. Carr (Queen Mary College) pointed out, it is not hard to think of explanations, and therefore Matsumoto's signal is of little help in distinguishing one cosmological model from another. The most prosaic origin is that the 1-mm excess is a redshifted remnant of the galactic 20-K emission which is prominent today; a population of early galaxies at high redshift is not difficult to find.

Slightly more exotic is the notion that a pregalactic generation of extremely massive stars, evolving on a very short-time scale, would leave behind a suitable radiation signal. In both these cases, care must be taken to ensure that the proposed source does not exceed observed emission levels at other wavelengths, but this constraint is neatly avoided in the most exotic explanation, espoused by J. Ostriker: a population of superconducting strings, whose main purpose is to initiate galaxy formation, produces radiation of a characteristic energy. The string population and its redshift can be chosen to give Matsumoto's infrared excess, and there is no significant signal elsewhere. David Lindley

*Comets to Cosmology, 3rd International IRAS Conference, London, 6-10 July 1987.

display; but one has the uncomfortable feeling that it simply is courtship display, and that its anti-jumping spider effect is no more than a happy accident.

Of course the display might have both functions, and if jumping spiders do exert strong predation pressure in the wild, as Green *et al.* assert, then there is every reason for courtship displays to take on the dual-purpose role. If it can be inferred that this behaviour is genuinely protective, then it does seem to be mimicry of a new kind. The converse type, where the predator mimics its prey, is well known; perhaps the best example is the firefly *Photurus versicolor*, where the female lures males of other species to their fate by imitating the flash patterns of their respective females (Lloyd, J.E. *Scient. Am.* 245(1), 138-145; 1981). Here we have the reverse, a prey species that gains by

imitating its predator.

The effectiveness of leg mimicry provides convincing, if circumstantial, evidence that leg patterns are indeed the main sign stimuli used by the jumping spiders to distinguish prey from conspecifics. I have made an observation that further supports this view, which is that on occasion jumping spiders mimic themselves. The figure shows a male of *Metaphidippus imperialis*, a species inhabiting Californian redwood forests where there is too little light for the black legs themselves to be seen properly. Presumably, the leg-like pattern on the face has the same function as the tetrithids' wing patterns: to convince other jumping spiders that this is indeed one of their own kind. □

M.F. Land is Professor of Neurobiology in the School of Biological Sciences, University of Sussex, Brighton BN1 9QG, UK.

Vision research

New views from compound eyes

Simon Laughlin

STUDIES of the compound eyes of arthropods are at the forefront of vision research, particularly in optics, phototransduction, colour vision, motion perception and pattern recognition. The arthropod visual system is also turning out to have properties allowing the evolution of vision to be elucidated. The optimistic mood at a recent meeting* suggests that the stage is set for a new era in research that integrates genome, structure, function and phylogeny at all levels, from molecule to behaviour.

Sigmund Exner 100 years ago discovered the lens cylinder in insect eyes and founded the field of graded refractive-index (GRIN) optics. He described two types of eye, apposition and superposition (see figure). Compound eyes continue to reveal new optical principles. In the apposition eyes of butterflies the tip of the cone is modified to act as a mode coupler. A refractive-index gradient in the cone tip matches the diffraction pattern of light from the lens to the diffraction pattern of the waveguide (the waveguide modes). This coupling improves both the sensitivity and the resolving power of the receptors (J.H. van Hateren, State University, Groningen). The refractive-index gradient in the cone tip is of the type required to make the lens cylinders of superposition eyes (see figure legend). Thus, a modification that improves the performance of the apposition eye of butterflies allows for the otherwise difficult evolutionary transition to the superposition eyes of moths (D.-E. Nilsson, University of Lund). There are also potential evolutionary pathways in crustaceans, where larval apposition eyes develop into adult superposition eyes via functional intermediates.

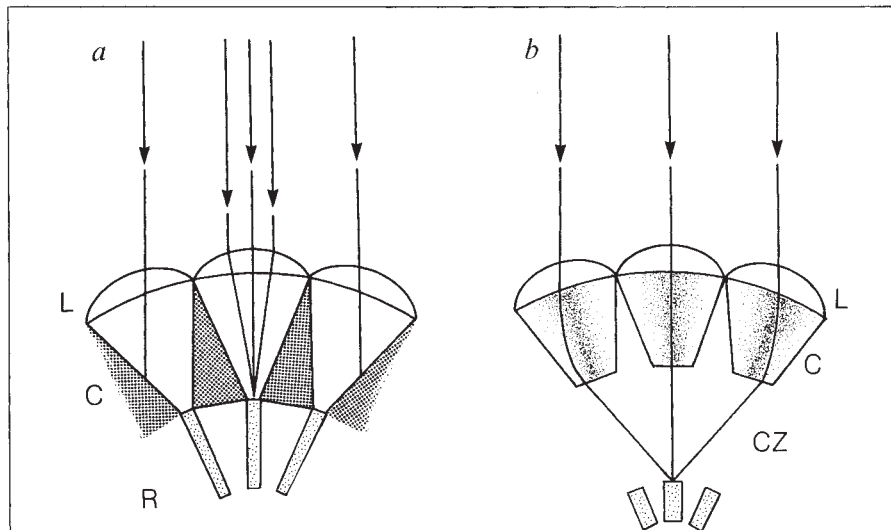
The sensation of light derives from the absorption of photons by the visual pigment molecules of the photoreceptor membrane, which consist of two parts, a small chromophore (retinal) held in a large protein (opsin). The protein is an enzyme that is activated by the structural changes occurring when the chromophore absorbs light. Many insect species use an unusual chromophore, 3-hydroxyretinal. The evolutionary perspective is exemplified by a systematic survey of chromophore distribution among 52 insect species from 15 orders (K. Vogt, University of Freiburg). Vogt suggests that the 3-hydroxy chromophore evolved when maggot-like larvae started to feed on rotting vegetation where xanthins, the

precursors of retinal, are scarce. Once evolved, this chromophore is phylogenetically stable, and its pattern of occurrence has implications for insect taxonomy. This new chromophore, originally a dietary imperative, may be a pre-adaptation for a refinement of insect vision — namely antennal pigments. Thus, the alcohol derivative (3-hydroxyretinal) is used by higher flies as an additional, antennal chromophore, which absorbs ultraviolet photons and transfers their energy to the chromophore proper. Sequence data for opsin are accumulating, and these may reveal the amino-acid substitution needed to accommodate new chromophores and sensitizing pigments.

The absorption of a photon by a visual pigment molecule sets in train a set of reactions culminating in an electrical response from the photoreceptor. The biochemical amplifier, coupling photopigment activation to current flow, was extensively discussed. Provocative evidence from the horseshoe crab *Limulus* suggests that light triggers formation of the second messenger inositol 1,4,5 trisphosphate, which stimulates the release of calcium from intracellular stores, generating the photocurrent by activating ionic

channels in the photoreceptor membrane (R. Payne, Marine Biological Laboratory, Woods Hole). Ironically, this second-messenger system, which resembles many other receptor-mediated responses, resurrects the calcium hypothesis, recently disproved for vertebrate photoreceptors. However, calcium-chelating agents, which block the response to injected inositol trisphosphate, only partially block the light response. There may be a second pathway involved, perhaps using the second messenger of the vertebrate photoreceptor, cyclic GMP (see, for example, Johnstone, E.C. *et al. Nature* 324, 468; 1986). There is evidence that fly photoreceptors also use an inositol trisphosphate system (B. Minke, Hebrew University, Jerusalem) and, with the fruitfly *Drosophila*, there is a formidable battery of genetic techniques available to study the regulation, development and the evolutionary history of this fundamental biochemical system.

In the first optic ganglion of flies, the photoreceptor signals are transmitted to interneurons across chemical synapses. The electrical response to light by the photoreceptor stimulates the release of synaptic-transmitter molecules which diffuse across the synaptic cleft to activate the interneuron. Fly photoreceptor synapses seem to use histamine (R.C. Hardie, University of Cambridge), which means they could be a good model for the study of the receptors and ion channels activated by this unusual neurotransmitter.



Compound eyes are arrays of optical subunits, shown here in diagrammatic cross-section. *a*, Apposition type; *b*, superposition type. Each unit consists of a facet lens, L (usually 20–30 μm in diameter), cone (C), and a column of photoreceptors (R). In the apposition eye the photoreceptor column acts as a waveguide of diameter 2–5 μm , with its receiving end apposed to the tip of the cone. Because the end of this waveguide lies in the focal plane of the lens, the receptor column receives light from one point on the image. Dense pigment between the cones ensures that each optical subunit operates independently. In superposition eyes, the photoreceptors and the cones are separated by a wide, clear zone (CZ). Each cone contains a well-defined refractive-index gradient, with the index falling radially from the cone axis. As Exner discovered, this refractive-index gradient caused the cones to act as lens cylinders, bending rays so that parallel light entering many facets is directed to one photoreceptor column. This superposition of light from several facets increases the sensitivity of the eye, hence its widespread occurrence in insects and crustaceans that are active in dim light.

* *Facets of Vision, Compound Eyes from Exner to Autrum and Beyond*, University of Regensburg, 21–24 April 1987. Organized by D. Burkhardt, R.C. Hardie and D.G. Stavenga.

The neuroanatomy of the first optic ganglion is known in great detail, providing a basis for the study of the evolution of neural circuits in dipteran flies (S.R. Shaw, Dalhousie University, Halifax). Progressing from primitive to advanced groups, the number of postsynaptic cells at each photoreceptor synapse increases from two to four, thus generating novel circuits (see the News and Views article by Melody Siegler, *Nature* 326, 130; 1987).

Once visual information is coded in neurons, what is done with it? In flies a small group of identified cells code motion of the retinal image: each cell responds to a unique combination of the direction of motion and of the size and position of moving objects. Microstimulation of pairs of identified photoreceptors generates vigorous movement signals in these neurons, and can elucidate fundamental spatial and temporal interactions underlying the computation of a motion signal (N. Franceschini, CNRS, Marseille). The motion-sensitive cells fall into two main classes, one best stimulated by wide-field motion, and the other preferring movement of small targets against a background. Microlesions selectively disconnect one or other cell type, and the resultant behavioural deficits suggest that the widefield cells code global flow-fields induced by body motion. The other type of cell is involved in following the movement of smaller objects against a larger moving background (K. Hausen, Max Planck Institut, Tübingen). The motion-sensitive cells project via interneurons to motor systems in a consistent and comparatively simple way (N.J. Strausfeld, University of Arizona, Tucson). Thus the fly promises to reveal some of the principles by which sensory arrays are mapped onto motor neurons, so as to achieve a precise and efficient motor control.

Colour vision is being systematically studied in bees and wasps (R. Menzel, Free University, Berlin) by measuring the sensitivity of photoreceptors to the wavelength of light and relating it to an insect's capacity to discriminate between different colours. The wavelengths that are most effective at stimulating the ultraviolet, blue and green photoreceptors vary remarkably little between species, but modelling suggests that the observed differences can have surprisingly large effects on wavelength discrimination. These differences could be important for animals living in different habitats. Such systematic comparative studies will enable us to relate the molecular structure of photopigments to behaviour, habitat, and hence evolution.

The sensitivity of a photoreceptor to different wavelengths depends principally on the visual pigment. The opsin component of the pigment tunes the chromophore to respond to a particular spectral band. Tuning depends on the structure of

the opsin and hence its amino-acid sequence. A surprising but plausible suggestion, based partly on comparisons between opsin sequence data from insects and mammals, is that molecular constraints are as significant as ecological and optical factors in determining the evolution of photoreceptor wavelength sensitivity (T. Goldsmith, Yale University). Particular tunings of the chromophore would require changes in opsin structure that are incompatible with the dual role of the visual pigment as a receiver of light and as a membrane-bound enzyme.

An elegant study of vision in *Limulus* (R.B. Barlow, Syracuse University) captured the spirit of the meeting by integrating optical, molecular and neural mechanisms with behaviour. At night, *Limulus* uses its lateral eyes to locate its mate.

Sensitivity is controlled by efferent nerves running from a circadian clock in the brain to the eye. Efferent activity sets the system to a high-sensitivity, night-vision state, and a circulating hormone drives the system back to day state when neural activity stops. The efferent input increases the aperture of the optical system, primes membrane synthesis, raises the gain and signal-to-noise ratio of the photoreceptors and tunes neural circuits. Behavioural experiments should reveal how these changes influence sensitivity and acuity. Because these changes resemble dark adaptation in many eyes, the results will be of general significance to the understanding of night vision. □

Simon Laughlin is in the Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK.

Astronomy

SS433 continues to perplex

Bruce Margon

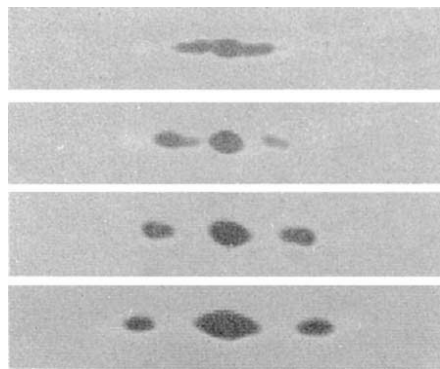
NEARLY a decade has passed since the popular press referred to the newly discovered object SS433 as the star that is "coming and going at the same time". Considerable observational and theoretical effort has been invested in this unique system¹, which ejects two opposed, highly collimated jets of gas at a remarkable 26 per cent of the speed of light. Yet if workers in the field were now to confront the press with a status report, the result would be a bit embarrassing. The problems posed by SS433 are difficult indeed and despite the intensive effort invested, the latest fruits of which are reported by Vermeulen *et al.* on page 309 of this issue², we are in many ways no wiser today.

Our ignorance is not the result of a lack of diligent or ingenious observations. The object is relatively bright at radio, infrared, visible and X-ray wavelengths. The basic nature of the jets was first identified in optical observations that also revealed the rotation of the jet axis once every 164

days. These results led to the most popular model of the system, in which the jets originate at, or near, a neutron star or black hole orbiting a massive and luminous, but otherwise normal, early-type star. At a recent meeting in St Louis, however, it became clear that the exciting results on SS433 are now being obtained using radio and X-ray observations. The new radio data reported in this issue² by Vermeulen and colleagues were obtained using elaborate very long baseline interferometry (VLBI) based on five radio-telescopes stretching from the United Kingdom to Italy. The authors achieve a resolution of about 10^{-2} arc seconds, corresponding to detail the size of the Solar System (100 astronomical units) at the distance of SS433 (5,000 parsec). This kind of resolution, or better, is essential to elucidate the physical conditions in the jet.

Peculiar emission lines in the visible spectrum gave the first indications that the jets consist chiefly of discrete 'blobs' of material rather than a continuous stream. The new VLBI data confirm this vividly and elegantly (see figure). They show non-thermal radio emission from the individual blobs as they coast ballistically away from the central cannon, along the path identified by the visible spectroscopy observations. We know which way this gun is pointed each day, but how and why it fires is still obscure. Vermeulen *et al.*² observed that the intensity of the radio emission from the blobs varies, consistently reaching maximum brightness about 6 days after launch. This is perhaps when they pass through the copious stellar wind thought to be lost from the companion star.

Recent X-ray observations from



Radio images of blobs ejected from SS433, taken at 2-day intervals by the European VLBI network. The pictures are approximately 700 astronomical units across.

EXOSAT reveal much about the origin of the bright point source associated with SS433. A strong line from highly ionized iron³ exhibits the same Doppler shift, and change in shift, as the visible emission lines. This implies that the X-ray emission is largely thermal and originates from the jets. Also, the accretion disk obscures one of the two jets at all times, so that the X-ray-emitting part of the jet is no longer than 10^{12} cm. As W50, the radio nebula that surrounds SS433, is elongated along the jet axis and can be mapped to a distance of 10^{20} cm from the central source, the jet can now be traced over about eight orders of magnitude of distance.

These latest X-ray results complement the earlier X-ray imaging observations of the Einstein Observatory that showed that the 10 per cent of the X-ray flux that is not associated with the central point source is aligned as a collimated jet, parallel to the W50 major axis. But the iron emission line from the core seen so prominently in the EXOSAT data was not reported by the earlier observers, who argued that the central X-ray source is not thermal. Now there can be little doubt that the X-ray emission from the centre of the SS433 is thermal emission from hot gas. This interpretation imposes a new and severe constraint for theories of the acceleration mechanism of the jet. It has been suggested^{4,5} that the jet has the particular observed velocity because of radiative acceleration from 'line locking', a fortuitous coincidence between the energy of ultraviolet photons from the central source and resonance transitions of abundant elements in the jet. This mechanism appears to be promising in principle, although rather complicated. But the iron X-ray lines are Doppler shifted, with a velocity 0.26c, implying that the jet achieves its terminal velocity after travelling only 10^{12} cm. The line-locking mechanism requires a much greater distance to deposit the requisite energy.

A reported detection of Doppler-shifted γ -ray emission⁶ from SS433 caused considerable excitement for several years; the implied luminosity was staggering, even by the superlative standards of SS433. There have since been several observations that failed to detect these fluxes, including data covering many phases and states of activity of the system. Perhaps the γ -ray results may now be removed from the long list of unsolved and tantalizing problems from SS433. □

1. Margon, B. *A. Rev. Astr. Astrophys.* **22**, 507-536 (1984).
2. Vermeulen, R.C., Scilizzi, R.T., Icke, V., Fejes, I. & Spencer, R.E. *Nature*, **328**, 309-313 (1987).
3. Watson, M.G., Stewart, G.C., Brinkman, W. & King, A.R. *Mon. Not. R. astr. Soc.* **222**, 261-271 (1986).
4. Milgrom, M. *Astr. Astrophys.* **78**, L9-L12 (1979).
5. Shapiro, P.R., Milgrom, M. & Rees, M.J. *Astrophys. J. Suppl. Ser.* **60**, 393-431 (1986).
6. Lamb, R.C. *et al. Nature* **305**, 37-39 (1983).

Bruce Margon is in the Department of Astronomy, FM-20, University of Washington, Seattle, Washington 98195, USA.

Mathematical models

Shell patterns simulated

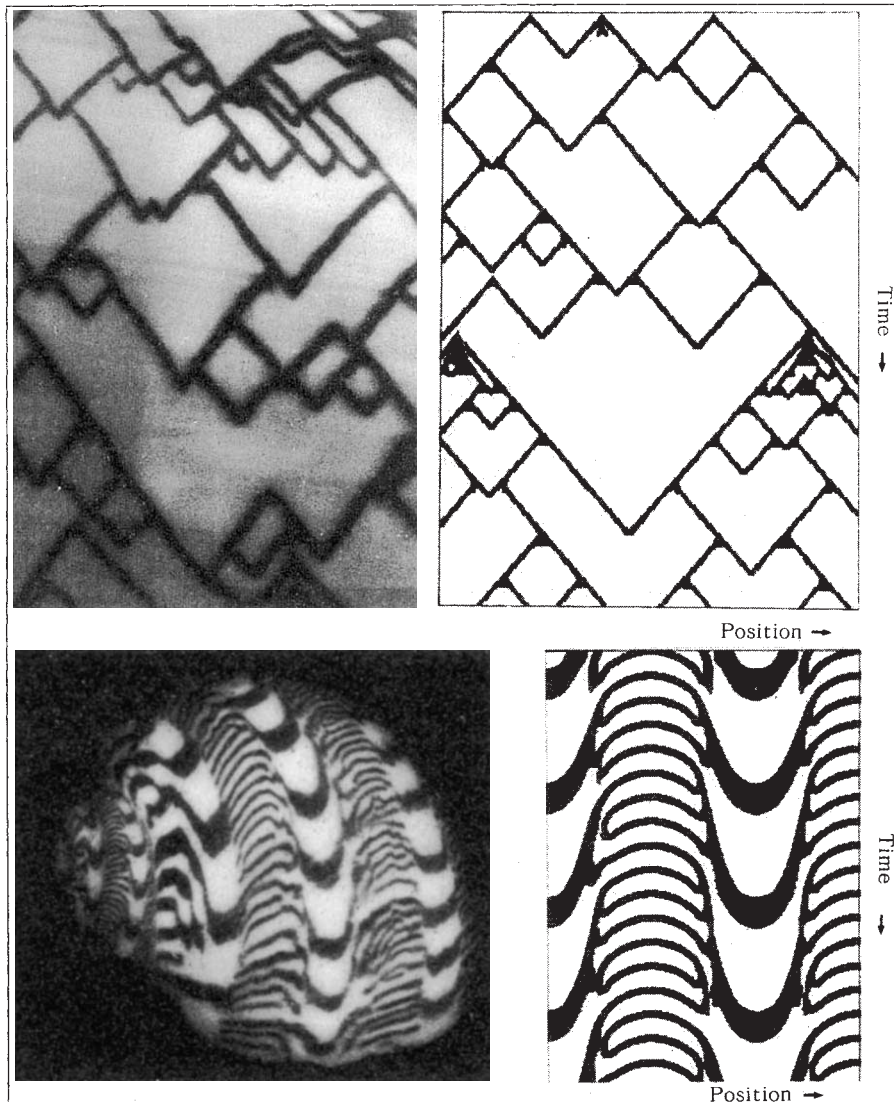
Geoffrey North

THE extraordinary variety of patterns exhibited by the shells of molluscs provide a good illustration of the combination of complexity and regularity that gives much of the beauty to the world of biology. A paper just published by Hans Meinhardt and Martin Klingler of the Max-Planck Institut für Entwicklungsbiologie in Tübingen (*J. theor. Biol.* **126**, 63-89; 1987) demonstrates that a relatively simple basic model can, with suitable adjustments, simulate many of the diverse shell patterns that occur. The picture shows just two of the patterns considered by Meinhardt and Klingler, with the simulations on the right of each pair closely matching the natural shell patterns on the left.

In the model, which is of the reaction-diffusion variety, the shell pigmentation patterns are considered to be the time records of the behaviour of a line of cells

constituting the growing edge of the shell, which forms the position dimension along which diffusion of the reacting chemical protagonists can take place. One of the protagonists is a diffusible 'activator' postulated to control pigment deposition, and it is the spatially and temporally varying level of this activator that generates the patterns. The synthesis of the activator is regulated by autocatalysis opposed by either substrate depletion or a simultaneously produced inhibitor. These reactions are simulated by coupled non-linear partial differential equations, the key parameters of which are the time constants of the reactions and diffusion rates of the activator and substrate or inhibitor.

Meinhardt has previously reported that such a model can simulate many of the simpler shell patterns, such as stripes of pigmentation perpendicular to the grow-



ing edge (spatial periodicity) or parallel to the growing edge (temporal periodicity), and even oblique lines resulting from travelling waves of pigment deposition. In their new work Meinhardt and Klingler show how the same basic model can be elaborated to generate some of the more complex shell patterns that are found. The top pair of pictures shows a detail of a shell of *Olivia porphyria* with oblique lines of pigmentation that occasionally branch and that extinguish where they collide. These two processes occur at a rate that maintains a roughly constant density of lines, an observation that provided a clue as to how the pattern could be generated. The extinction of travelling waves is a natural product of the model, resulting from the refractory state that cells pass through in the wake of the waves. To generate branch points and thus new lines an additional factor is needed: Meinhardt and Klingler postulate a rapidly diffusible 'hormone-like' regulator that controls the rate of inhibitor decay, and the production of which is a function of the number of waves. As waves collide and extinguish the level of the hormone regulator de-

creases, increasing the rate of inhibitor decay and as a consequence promoting branch formation.

Natica enzona, the shell below on the left, illustrates the spatial variation in temporal oscillation frequency sometimes observed; bands with a low oscillation frequency alternate with bands of higher oscillation frequency. This is simulated by the superimposition of the kind of pattern-generating system used to generate time-invariant stripes perpendicular to the growing edge with a time-varying oscillation in pigment production: along the edge there are two alternating levels of substrate concentration that cause either a high or low frequency of oscillation.

Those who imagine that a discrete model might be more realistic should perhaps urge their nearest cellular automata theorist to try their hand at matching (or improving on) Meinhardt and Klingler's simulations (see Wolfram, S. *Nature* **311**, 419-424; 1984). There are still some patterns Meinhardt and Klingler have not been able to simulate. □

Geoffrey North is Deputy Biology Editor of *Nature*.

Materials science

Permeability of nuclear fuel

Robert W. Cahn

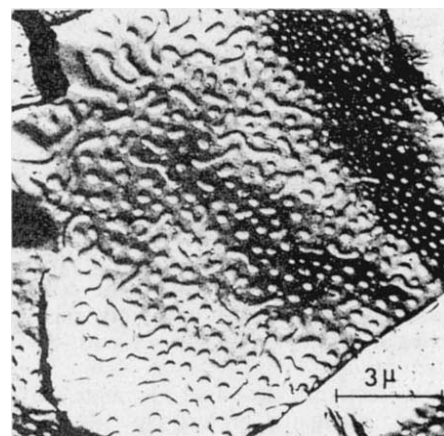
CERAMIC objects, including the new high-temperature superconducting ceramics, are generally made by sintering a compacted mass of powder. In the process, empty volume gradually diminishes and the compact shrinks. When the residual porosity has decreased to 5-6 per cent, it changes from a connected network of channels to a system of isolated pores, and the permeability to gas of the compact disappears. After about 40 years of microstructural experiments, the mechanisms of sintering are now reasonably well understood¹. In a different context, ceramic fuel in a nuclear reactor — generally uranium oxide, UO_2 — gradually accumulates bubbles filled with fission-product gases, krypton and xenon, at high pressure (see the figure). These bubbles can change from an isolated, non-permeable configuration to a permeable network of communicating channels along the grain edges — the reverse of sintering — and this transition for an equiaxed grain structure again comes at 5-6 per cent. A recent paper² by Nichols brings our understanding of this obviously important subject a large step forward.

This type of deterioration of nuclear fuels has been thoroughly investigated, because whereas isolated bubbles keep the gas safely ensconced in the ceramic, permeable fuel releases the radioactive gas to form a high-pressure reservoir that

might fracture the protective fuel can. In turn, this would release radioactivity to the coolant. Of course, only operating conditions that do not lead to the growth of a continuous network of channels are used in nuclear reactors. The more difficult it is to form such a network, the longer a fuel rod can safely be left in the reactor.

In geology also, as it happens, the conditions favouring connected channels of liquid in partially molten rock are of interest, and have been analysed recently³. The principal concern here, however, is with the kinetics of the expulsion of liquid rather than with identification of equilibrium configurations.

There are many models of the critical porosity observed to bring on communicating networks of channels. One approach uses stereological analysis of the topology of a powder during sintering. The crucial variable here is the connectivity, G , where $G = C - P + 1$, C is the total number of interparticle contacts and P is the number of particles. Rhines has shown how the connectivity can be measured⁴, and that it diminishes during sintering¹. From this, and with further stereological characterization, the sintering force (the critical applied tensile force needed to arrest the contraction resulting from sintering) can be calculated and compared with experiment. For the early stages of sintering, calculation and measurement



Early stages of bubble linkage at a fractured grain boundary in UO_2 , with incipient channel formation at the bottom right. (By Hyam, E., from p. 422 of ref. 8.)

are in good agreement, but as porosity declines below 20 per cent, there is great disparity. Plainly, microstructural analysis does not suffice to interpret the critical stage at which pores become isolated.

Another approach for the inverse transition from communicating to isolated porosity, is to analyse the stability of grain-edge channels in terms of variables such as specific surface and grain-boundary energies. This approach has now been greatly simplified in the recent paper by Nichols². His treatment is based on a famous mathematical analysis by Lord Rayleigh⁵ — possibly the oldest frequently cited reference in the whole of materials science — of the instability criterion for an inviscid fluid jet as the jet's dimensions are varied. If the jet length λ exceeds a critical value $\lambda_c = 2\pi R$ (R is the jet radius), then an infinitesimal shape perturbation will be unstable so that the jet separates into distinct drops. Nichols has applied this kind of analysis to an elongated channel attached to a grain edge: if $\lambda \gg \lambda_c$, the channel will break up into shorter lengths λ_m , where λ_m varies between 9 and 13 R according to the dominant mass-transport mechanism. In Lord Rayleigh's analysis, the length of the drops resulting from the break up of a long jet is 9 R .

The applicability of Rayleigh's analysis has been tested by McLean^{6,7}, using transmission electron microscopy, with a hot stage, of thin aluminium foils containing a dispersion of long molten lead rods up to 1 μm in diameter. Drop formation begins preferentially at the rod ends (at a rate varying as R^{-2}). McLean points out that Rayleigh had not predicted the preferential end-effect, but then Rayleigh's analysis applied to effectively infinite jets, just as Nichols' analysis applies to effectively endless channel networks.

In the operation of a cylindrical nuclear-fuel element, the UO_2 grains gradually become elongated in the direction of the (radial) thermal gradient. So Nichols analysed the behaviour of a channel net-

work in a regular array of parallel, square-based rectangular parallelepipeds as b/a (height divided by base) varies. The volume fraction of porosity is calculated as a function of R , a and b , and the Rayleigh-type analysis predicts instability for $b = 2\lambda_c = 4\pi R$. The analysis shows that for $b = a$ (an equiaxed grain) the critical porosity to permit a stable channel network is 5.2–5.5 per cent, in accord with experiment. But as b/a increases, the critical porosity for channel instability rapidly rises to 10 per cent for $b/a = 2$, and 40 per cent for $b/a = 4$.

In fact, practical fuel elements, with highly elongated grains, typically show porosities of only a very few per cent, and this can be so only if no channel network exists. The conclusion is that the very

pronounced grain elongation ($b \gg a$), that always accompanies the burn up of highly rated fuel is in fact an excellent natural safeguard for the integrity of the fuel, preventing the release of radioactive fission gases. Nature, it appears, does not have to be green to be benevolent. □

1. Rhines, F.N. *Microstructure — Behaviour and Microstructure of Materials*, 49–59 (Riederer, Stuttgart, 1986).
2. Nichols, F.A. *J. nucl. Mater.* **148**, 92–98 (1987).
3. McKenzie, D. *Earth planet. Sci. Lett.* **74**, 81–91 (1985).
4. Rhines, F.N. & Chang, S.S. *Scr. Metall.* **18**, 371–372 (1984).
5. Lord Rayleigh, *Proc. Lond. math. Soc.* **10**, 4–13 (1878).
6. McLean, M. *Phil. Mag.* **27**, 1253–1266 (1973).
7. McLean, M. & Loveday, M.S. *J. mater. Sci.* **9**, 1104–1114 (1974).
8. Gittus, J. *Irradiation Effects in Crystalline Solids* (Applied Science, London, 1978).

Robert W. Cahn is at the Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK.

Superconducting ceramics

Applications just beyond reach

A.D. Appleton, A.M. Campbell and David Caplin

SUPERCONDUCTIVITY may seem to be a single topic of specialist interest, but a recent conference* was attended by solid-state physicists and chemists, materials scientists, and electrical engineers interested in microelectronics and power. This breadth of representation illustrates graphically the wide-ranging appeal of the new and exciting developments.

It was emphasized (M. Rice, ETH, Zurich) that transition-metal oxides are indecisive beasts, rarely content to remain as simple metals; instead, they are susceptible to spin-density waves, or charge-density waves, or some other form of electronic ordering. So, although some groups claim that undoped La_2CuO_4 is superconducting, others suggest that superconductivity occurs only if the sample is slightly off-stoichiometry, or contains impurities. $\text{YBa}_2\text{Cu}_3\text{O}_7$ appears to be an extraordinarily robust compound; by now it must have been made by hundreds of research groups, and all are agreed that it becomes superconducting within a degree or two of 92 K. But in its normal state, is it a simple metal? The normal-state resistivity is sufficiently high (Rice), approaching 10^{-3} ohm cm, that electron-phonon scattering is unlikely to be responsible for the observed linear dependence of the electrical resistivity on temperature.

Careful and systematic work (K. A. Muller, IBM, Zurich) has shown that both the $(\text{La,Sr})_2\text{CuO}_4$ and $\text{YBa}_2\text{Cu}_3\text{O}_7$ compounds have jumps in their specific heat capacity at the superconducting transition temperature. In sharp contrast to the behaviour of conventional superconductors, however, his data suggest that in the new

materials, a linear term in the electronic heat capacity survives to low temperatures. Although other mechanisms might be involved, this could imply that there are two mechanisms that co-exist, one with a gapless spectrum of electronic states and the other with the gap seen by tunnelling and other experiments.

The role of oxygen in the lattice is of vital importance and the details remain to be established. It appears that the highest transition temperature is in stoichiometric $\text{YBa}_2\text{Cu}_3\text{O}_7$, but that ordering of oxygen vacancies where they exist may be important. Perfect single crystals were not reported, but small crystals with parallel c -axes twinned in the basal planes were (G. Balistrino, Rome). As the difference between the a - and b -axes is small, these are good approximations to single crystals.

All the bulk samples contain grain boundaries that are relatively wide and amorphous. It is not clear at present whether this is because of impurities or because of disordering of the superconducting phase itself, but it has a crucial effect on the superconducting properties and is the main obstacle to the use of these materials.

In many ways it is not the high critical temperatures of the new materials, but their high critical fields (which quench superconductivity) that are their most spectacular property. At 9 tesla (T) the critical temperature for the perpendicular orientation was reported to be 87 K (H. Takagi, Tokyo), although an anisotropy of a factor of 5 means lower critical fields parallel to the base plane and thus it will be necessary to orientate the material to make the most of its properties. At 95 T, the magnetic energy is 4×10^9 J m $^{-3}$, corresponding to a pressure of 4,000 MPa.

Although the technical problems are formidable, these figures illustrate some of the potential of these materials.

The achievement of high critical-current densities remains the obstacle to the use of new materials. Magnetization measurements show that these are conventional type-II superconductors to which earlier theories can be applied, and therefore the problem is one of sample preparation. The material should be able to carry 10^4 A cm $^{-2}$ in 10 T. Current densities up to 10^6 A cm $^{-2}$ within the grains were measured at 4.2 K, but insulating or amorphous layers between the grains prevent the current passing between them. This problem also occurred in thin films (A.M.C.) and appears to be more acute in oxide superconductors than in traditional intermetallic compounds. There are several promising processing routes that should break up these layers (such as hot isostatic pressing) but much work is needed to get the conditions right. Considerable progress has been made in the production of multifilamentary wires, but again contact between grains is poor and, so far, current densities are low.

Rather surprisingly it has been found that lumps of oxide behave as superconducting quantum interference devices (SQUIDs) that are used for sensitive magnetic-field measurements, without any special processing (see Colclough, M.S. *et al.* *Nature* **328**, 47–48; 1987). Calculations show that they should have a similar performance to liquid-helium-temperature SQUIDs (C. M. Pegrum, Strathclyde Univ). For better control, and microcircuit applications, it will be necessary to produce thin films and one group (M. G. Blamire, Cambridge) was able to show sputtered films, or material, etched into device structures. The oxide barrier was too thick to show tunnelling characteristics and work is continuing in order to produce narrower resistive transitions and thinner barriers.

The large-scale applications for high-temperature superconductors depend upon the availability of wire or ribbon that can carry large currents and be wound into coils. If it is assumed that these materials are available it is possible to achieve some notable improvements in the marketability of large-scale devices. The key to this is that industry is used to handling liquid nitrogen but not liquid helium. In general, the cost of motors, generators, cables, levitated transport, fault-current limiters and so on will be reduced but the greater reliability that results from simplified designs will be the spur to greater investment. □

A.D. Appleton is at NEI International Research and Development, Fossway, Newcastle upon Tyne NE6 2YD; A.M. Campbell is in the Department of Engineering, Cambridge University, Cambridge CB2 1PZ; and David Caplin is at the Department of Physics, Imperial College, London SW7 2BZ, UK.

* European Workshop on High- T_c Superconductors and Potential Applications, Genoa, 1–3 July 1987.

Neurophysiology

Dual roles for DHP receptors in excitation-contraction coupling?

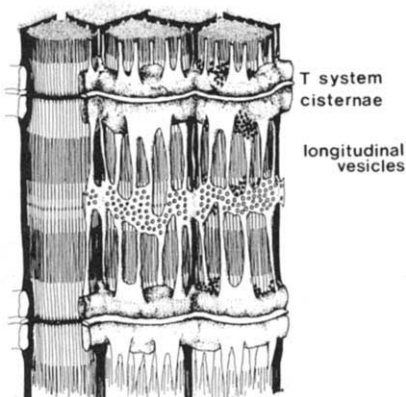
William S. Agnew

THE report on page 313 of this issue by Tanabe *et al.*¹ in S. Numa's laboratory of the molecular cloning of the skeletal muscle dihydropyridine (DHP) receptor promises to shed new light on the roles this protein may serve as a voltage-activated Ca^{2+} channel and as a voltage-sensor linked to excitation-contraction coupling.

Dihydropyridines are a well-studied class of Ca^{2+} channel antagonists that have been used as biochemical markers for receptor purification. DHP receptors isolated by several groups comprise a large peptide of relative molecular mass 135,000–175,000 (135–175K) and one or two smaller subunits of 32–34K and 50K (see refs 2, 3, but also ref. 4). Numa's group¹ has now cloned and sequenced the complementary DNA encoding the large subunit, with intriguing results. As discussed by Stevens in his News and Views article⁵ last week, the predicted sequence of 1,873 amino acids (212K) shows fundamental similarities with previously cloned voltage-sensitive sodium-channel peptides⁶. Among the similarities are four internal repeats each of about 25K, which might fold into membrane-spanning pseudosubunits (see cover of last week's issue). Within each repeat are sequences closely similar to the distinctive S4 segments thought to act as voltage-sensors in sodium channels. Taken together, the sequence data strongly suggest that this peptide forms a voltage-sensing molecule, probably an ionic (Ca^{2+} selective?) channel. The presence of a long hydrophilic (presumably cytoplasmic) carboxy terminus bearing six consensus phosphorylation sites suggests that one or more of its functions are closely regulated.

What does this structure imply about the roles of the molecule in excitation and contraction? It is well known that T tubules extend from the sarcolemma of muscle fibres to the cell interior, where they come into apposition with cisternae of the sarcoplasmic reticulum that contain intracellular Ca^{2+} stores (see figure). Bridging the membrane triad (a central T tubule flanked on either side by sarcoplasmic reticulum) is a regular array of 120–200-Å particles, or 'feet'. Electrical impulses spreading from the sarcolemma into the T-tubule network somehow stimulate the sarcoplasmic reticulum to release Ca^{2+} and initiate muscle contraction. As Somlyo has discussed⁸ in News and Views, there are several hypotheses that could account for excitation-contraction

coupling. According to one, DHP-sensitive T-tubule Ca^{2+} channels produce localized Ca^{2+} currents which trigger Ca^{2+} -dependent Ca^{2+} release from the sarcoplasmic reticulum. Alternatively, a DHP-sensitive voltage-sensor in the T tubule might activate a second-messenger system, possibly a phospholipase that would provoke the highly localized production of inositol trisphosphate, a compound capable of releasing Ca^{2+} from the sarcoplasmic reticulum stores. A third suggestion is that the voltage-sensors act via a conformational mechanism through the



Arrangement of the sarcoplasmic reticulum around myofibrils in vertebrate skeletal muscle (from Squire, J., *The Structural Basis of Muscular Contraction*, Plenum, 1981).

feet to activate Ca^{2+} release sites mechanically. Assuming that there are not two DHP-sensitive voltage-regulated proteins in T tubules, DHP receptors appear to have split personalities, capable of acting both as voltage-sensors and as voltage-activated Ca^{2+} channels in excitation-contraction coupling.

Before the report in this issue, experimental evidence has always fallen just short of equating DHP receptors with Ca^{2+} channels. Skeletal muscle, for example, has two Ca^{2+} currents, I_{fast} and I_{slow} . The latter appears to be carried by L-type Ca^{2+} channels and is blocked by dihydropyridines⁹. In skeletal muscle of mice that have genetic muscular dysgenesis, I_{slow} is missing¹⁰ and DHP receptor levels are greatly reduced. The initially perplexing finding that DHP blockade of Ca^{2+} currents requires 10–100-fold higher concentrations than does binding to receptor sites in membrane fragments can be explained by the discovery that depolarization greatly increases DHP-binding affinities¹¹.

Paradoxically, Ca^{2+} currents seem not to be required for excitation-contraction

coupling; they are slower and last longer than activation of contraction, and their elimination, by removing extracellular Ca^{2+} or by blocking with Cd^{2+} , does not prevent contraction⁸. Further, comparing Ca^{2+} current densities with the total number of DHP receptor sites indicates that, at most, only a small percentage of the molecules normally function as channels^{2,11}.

That DHP receptors are voltage-sensors linked to excitation-contraction coupling is suggested by the fact that both muscle contraction and movement of a substantial membrane charge (charge 1) associated with activation of the sarcoplasmic reticulum are quite sensitive to dihydropyridines⁷. Also, the muscles of dysgenic mice that have greatly reduced DHP receptor levels and slow Ca^{2+} currents fail at the excitation-contraction coupling step^{8,10}. The possibility thus arises that the DHP receptor may serve two (or more) independent functions.

Numa and colleagues in their paper in this issue propose that Ca^{2+} currents are small because they serve only to replenish intracellular stores. Indeed, a protein phosphorylation regulatory system could maintain Ca^{2+} currents at the low levels sufficient to meet the physiological requirements. Processes that trigger Ca^{2+} release from the sarcoplasmic reticulum could be coupled independently to voltage-dependent conformational changes, regardless of whether the ionic pathway was opened. These could include activation of the phosphoinositide system, or conformational coupling via the feet. We can expect to hear soon whether the cloned DHP receptor, expressed alone or with other peptides, can function as a Ca^{2+} channel. Further, the single recessive gene defect in murine muscle dysgenesis could affect the structure or regulation of the DHP receptor. If cultured dysgenic cells could be mended by introduction of the cloned gene, site-directed mutagenesis could then be used to pry open the secrets of excitation-contraction coupling. One only wishes that the clones described here¹ would become freely available to the many laboratories interested in these questions. □

1. Tanabe, T. *et al.* *Nature* **328**, 313–318 (1987).
2. Curtis, B.M. & Catterall, W.A. *Biochemistry* **25**, 3077–3083 (1986).
3. Schmid, A., Barhanin, J., Coppola, T., Borsotto, M. & Lazdunski, M. *Biochemistry* **25**, 3492–3495 (1986).
4. Leung, A.T., Imagawa, T. & Campbell, K.P. *J. biol. Chem.* **262**, 7943–7946 (1987).
5. Stevens, C.F. *Nature* **328**, 198–199 (1987).
6. Noda, M. *et al.* *Nature* **320**, 188–192 (1986).
7. Franzini-Armstrong, C. *J. Cell Biol.* **47**, 488–499 (1970).
8. Somlyo, A. *Nature* **316**, 298–299 (1985).
9. Rios, E. & Brum, G. *Nature* **325**, 717–720 (1987).
10. Beam, K.G., Knudsen, C.M. & Powell, J.A. *Nature* **320**, 168–170 (1986).
11. Schwartz, L.M., McCleskey, E.W. & Almers, W. *Nature* **314**, 747–751 (1985).

William S. Agnew is in the Department of Cellular and Molecular Physiology at Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510, USA.

Protein folding

A case of mistaken identity

Roger H. Pain

THE structure of molecular biology is built on several basic hypotheses, including the belief that one gene encodes one polypeptide chain and that the product of expression of a single gene is associated with a single biochemical function or activity. It has been realized for some time that the relation between DNA and protein-mediated functions is more complex, but the recent report¹ by Pihlajaniemi *et al.*, concluding that the product of a single gene seems to possess two entirely different enzymatic functions, one associated with the structure of organisms and the other with protein folding, challenges some of these intuitively held assumptions.

Elephants are largely held together by collagen, and the biosynthesis and assembly of this deceptively simple protein require the involvement of a dozen or more enzymes. In particular, the stability of the three-stranded helix of collagen depends on a proportion of the proline residues in the characteristic repeating sequence Gly-X-Pro being hydroxylated to hydroxyproline after synthesis of the collagen polypeptide. This reaction is catalysed by prolyl 4-hydroxylase, an enzyme comprising two α -subunits and two β -subunits. The two β -subunits are encoded by different genes and assemble from a large pool of β which is present in excess. The β -subunit is intimately involved in prolyl hydroxylase activity, because the α -subunit is inactive alone, because the whole enzyme is inactivated by an antibody specific for the β -subunit, and because both

subunits contribute to the ascorbate-binding site. As part of an extensive research programme on the biosynthesis of collagen, Pihlajaniemi and colleagues cloned and sequenced the gene encoding the β -subunit of prolyl 4-hydroxylase¹, thereby deriving the amino-acid sequence of the protein itself.

Apart from providing essential information for mechanistic and structural investigations, protein sequences are the basis for determining family relationships between proteins, and the existence of computerized databanks allows rapid comparison of new sequences with the many proteins whose sequences are already known. Operation of this now routine procedure on the sequence of the β -subunit threw up an entirely unexpected result: 94 per cent of the amino acid residues comprising the β -subunit of human prolyl 4-hydroxylase are identical to those in the rat enzyme, protein disulphide isomerase (PDI). Assaying the β -subunit shows that it has the same level of PDI activity as the enzyme previously isolated under this name. Further, the appropriate DNA probe shows the existence of only one gene encoding the common sequence in the human genome. In other words it seems highly likely that the small differences in sequence reflect species differences, and that the β -subunit of prolyl 4-hydroxylase is the same protein as PDI, both enzymic activities being encoded by a single gene.

Protein disulphide isomerase is an enzyme thought to catalyse the rate of reformation of correct disulphide bridges, hence allowing refolding of globular proteins whose three-dimensional conformation has been unfolded by urea and whose disulphide bridges have been broken by a reducing agent². Its tissue distribution has been studied extensively and its highest specific activities are in collagen-synthesizing cells. Its distribution correlates closely with that of prolyl 4-hydroxylase, although the two enzymes were at that time presumed to be different³. The sequence reveals two pairs of cysteine residues, each set in a sequence of 16 residues of nearly identical sequence and whose function is still uncertain but which, as in thioredoxin, could function as an oxidation-reduction couple interacting with cysteine residues in refolding proteins⁴. Although located in the cell and the endoplasmic reticulum, where protein folding takes place, the hypothesis that PDI is physiologically important in the folding of proteins *in vivo* is still unconfirmed.

What is the significance of this in-

triguing finding? Are there two distinct and unconnected enzyme activities, with two separate active sites on the same polypeptide chain together with binding sites for the α -subunit? If so, do these activities reside on separate — and separable — domains, suggesting an earlier gene-fusion event, using the possibly redundant PDI as a convenient piece of protein to stabilize the interaction of the complementary α - and β -catalytic units? Or has PDI been viewed in the wrong light all along? PDI in collagen-synthesizing tissue can now be seen to be prolyl 4-hydroxylase, whose function is well established, obviating the previously held view that PDI is required in collagen biosynthesis.

Could it be that those seeking the connection between the two enzyme activities should focus on their ability to bind proline? The rate-limiting step of protein folding is frequently the requirement for proline residues to take up one of two possible isomeric configurations, *cis* or *trans*, before the polypeptide chain can condense and pack optimally. This requirement applies to collagen as well as to many globular proteins⁵. Could the enzyme that binds proline in collagen in order to hydroxylate it also bind proline in a refolding polypeptide and thereby catalyse the *cis-trans* isomerization reaction?

The case would be argued first on the grounds that, for PDI to increase the rate of refolding, some reducing agent must be present to allow disulphide interchange to occur — only more quickly now that the accompanying non-covalent reshuffling is not being delayed by unaided proline isomerization. The hypothesis involves the interaction of 'PDI' with a proline residue at the surface of the collapsed and folding protein. This scheme is simpler to envisage than a model⁶ in which a disulphide bond on PDI has to interact in an analogous fashion with two separate sulphydryls. This alternative view would involve, however, a rather strange shifting of the *cis-trans* proline equilibrium on binding rather than classical catalysis, and to what extent this would result in a higher proportion of proline configurations that fit into the native protein conformation is not immediately obvious.

It is to be hoped that this latest discovery¹, serendipitous though it may be, will lead to further studies on this interesting molecule to clarify the details of how it performs these two seemingly unconnected functions. □

1. Pihlajaniemi, T. *et al.* *EMBO J.* **6**, 643–649 (1987).
2. Freedman, R.B. *Trends biochem. Sci.* **9**, 438–441 (1986).
3. Brockway, B.E., Forster, S.J. & Freedman R.B. *Biochem. J.* **191**, 873–876 (1980).
4. Edman, J.C. *et al.* *Nature* **317**, 267–270 (1985).
5. Freedman, R.B. *Nature* **279**, 756 (1979).
6. Creighton, T.E., Hillson, D.A. & Freedman, R.B. *J. molec. Biol.* **142**, 43–62 (1980).

Roger H. Pain is Professor of Physical Biochemistry at the University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK.



100 years ago

GEOGRAPHICAL NOTES

THE rumour as to the death of Mr Stanley is universally discredited in geographical circles. The rumour seems quite inconsistent with the news as to Mr Stanley's having left the Aruwimi River on June 3 for Wadelai. Had he been shot, as reported, it must have been after this date, and during the land journey, whereas one version of the rumour gives out that he was killed on the Congo. He may meet with Emin Pasha sooner than he expected. Emin, it seems, is at present exploring on the south of the Albert Nyanza so that he and Mr Stanley may meet half way. Letters from Mr Stanley are expected in this country early in August. From *Nature* **36**, 309; 28 July 1887.

Are plasmons the key to superconducting oxides?

SIR—Recently, Sir Nevill Mott addressed the issue of distinguishing viable explanations for the high-temperature superconductors¹. His discussion of polaron effects represents a novel approach to superconductivity in addition to the highly original suggestion proposed by Anderson on the basis of resonating valence bonds². Another approach to creating room-temperature superconductors relies on the plasma oscillations of electrons (or holes). In view of the reported experimental evidence for plasmon anomalies in the cuprates, it seems appropriate to indicate the limitations as well as the promising features of this electronic mechanism for superconductivity.

Traditional explanations of superconductivity based on phonon-mediated interactions (Bardeen-Cooper-Schrieffer (BCS) theory) provide the somewhat modernized formula for the superconducting transition temperature

$$T_c = 0.7 \theta \exp \left[-\frac{1+\lambda}{\lambda-\mu^*} \right]$$

where θ represents a typical temperature for the sound wave of the lattice vibrations, λ is the electron-phonon coupling and μ^* is the modified Coulomb repulsion between electrons. But in the case of phonon-mediated interactions, $\mu^* \approx 0.1$, $\lambda \approx 1$, and a value of $\theta \approx 300$ K yields superconducting temperatures up to 25 K. The case of Y-Ba-Cu-O has become more mysterious with the measurements where the ¹⁸O isotope is substituted for the normal ¹⁶O atoms. The Raman scattering shifts of the oxygen vibrations predict a change in T_c by 4 K according to the BCS formula, but the actual measurements rule out changes in excess of 0.2 K. In other words the evidence suggests minimal involvement by the oxygen-related phonons in the high-temperature superconductivity^{3,4}.

The original discovery⁵ of the La-Ba-Cu-O alloys showed that their T_c was unusually sensitive to the alloy content: for example, if 5% of the La is substituted by Ba, no benefit is produced, whereas 15% Ba triggers a superconducting state, with $T_c \approx 40$ K. Later, Chu and his colleagues showed that this can be improved by substituting yttrium for La, giving the Y-Ba-Cu-O series of 90 K superconductors⁶.

One interpretation of the concentration dependence of T_c is that a low frequency plasmon mode is created only in those alloys where the Fermi energy may intersect a narrow band of electrons (or holes), which then participate, in addition to the majority carriers. This suggestion⁷ is compatible with both recent electronic structure calculations for the Y-Ba-Cu-O compounds and photoemission data.

Moreover, the nature of the two component plasmons may resemble acoustic waves, with a corresponding 'sound' velocity much higher than ordinary phonon velocities.

The allure of plasmons as candidates for the mediation of the electron-electron pairing is obvious from the T_c formula. If the prefactor $\theta \sim 100,000$ K is substituted, the resulting T_c enhancement would indicate that room-temperature superconductors should be readily available. Unfortunately, this optimistic glance must contend with a disagreeable fact; namely, the existence of plasmons in metals is commonplace, yet ordinary metals are usually not very good superconductors, and some are not superconducting at all.

A plausible solution to this dilemma may be the existence of two plasmon modes: the first high-energy excitation, say $\Omega \sim 1$ eV, simply represents the majority of electrons or holes that carry electrical current, while a second branch at lower energies, say $\theta \sim 0.1$ eV, may occur under exceptional circumstances and meet the stringent requirements for an electron pairing interaction. The existence of an acoustic plasmon in metals was originally proposed by Pines, and applied

to superconductivity by Radhakrishnan⁸, Fröhlich and others⁹.

Infrared absorption experiments should provide a test of the plasmon models proposed to date. The two-component acoustic plasmons⁷ in the Cu-O plane should be very responsive to alloy composition and the interlayer plasmon¹⁰ should exhibit a selective response to polarized light. The single-component, two-dimensional plasmon invoked by Kresin¹¹ should be examined as well. The first evidence for a two-dimensional acoustic plasmon in a metal has just been reported for La₂NiO₄ (ref. 12).

JOHN RUVALDS

Lyman Laboratory of Physics,
Harvard University,
Cambridge,
Massachusetts 02138, USA

1. Mott, N.F. *Nature* **327**, 185 (1987).
2. Anderson, P.W. & Abrahams, E. *Nature* **327**, 363 (1987).
3. Batlogg, B. *et al.* *Phys. Rev. Lett.* **58**, 2333 (1987).
4. Bourne, L.C. *et al.* *Phys. Rev. Lett.* **58**, 2337 (1987).
5. Bednorz, T.G. & Müller, K.A. *Z. Phys.* **B64**, 189 (1986).
6. Wu, M.K. *et al.* *Phys. Rev. Lett.* **58**, 908 (1987).
7. Ruvalds, J. *Phys. Rev.* **B35**, 8869 (1987).
8. Radhakrishnan, V. *Phys. Lett.* **16**, 247 (1965).
9. Ruvalds, J. *Adv. Phys.* **30**, 677 (1981).
10. Ashkenazi, J., Kuper, C.G. & Tyk, R. preprint.
11. Kresin, V. *Phys. Rev. B* **35**, 8716 (1987).
12. Bassat, J.M., Odier, P. & Gervais, F. *Phys. Rev.* **B35**, 7126 (1987).

Extraordinary non-spreading light beams

SIR—The non-spreading light beam of Durnin *et al.*¹ mentioned in a recent News and Views article² is much like that produced by passing plane waves through a piece of glass flat on one side and conical on the other³. Both go on to form a circle at infinity. In getting there, the light passes through the axis, where it creates a region of high intensity. This region is a caustic, not a focus: in any transverse plane, only a minute fraction of the energy in the beam passes through the central bright spot. Because the caustic is straight, the conical lens and other 'axicons' (from axial icon = axial image) have been used as aligning devices. The device of Durnin *et al.* differs from every other axicon I know of in producing a diffraction pattern that ideally varies neither in form nor intensity with position along the caustic.

The caustic is limited in length to the portion of the axis crossed by the conical beam on its way to infinity. In the experiment of Durnin *et al.* this was about 0.8 m. It can be increased by narrowing the cone angle of the beam — but the central maximum will then be broader — or by increasing the diameter of the axicon. Fourier still rules the photon shooting gallery.

To hit a target with photons, there is still no better way than launching spherical waves, with the target as centre, from a spotlight of the largest possible diameter. Depending on how errors are penalized, it helps to vary the amplitude across the

launched wavefronts in one way or another. The root-mean-square miss distance is minimized if the amplitude falls off with distance from the axis as the J_0 Bessel function, scaled so its first zero falls at the edge of the beam (ref. 4 — in equation 4 of which the number 2.405 should be replaced by 2.405²). For uniform amplitude, the root-mean-square error is infinite in paraxial theory.

C.W. McCUTCHEN

Building 8, Room 403,
National Institutes of Health,
Bethesda, Maryland 20892, USA

1. Durin, J. *et al.* *Phys. Rev. Lett.* **58**, 1499 (1987).
2. Maddox, J. *Nature* **327**, 183 (1987).
3. McLeod, J.H. *J. opt. Soc. Am.* **44**, 592–597 (1954).
4. McCutchen, C.W. *J. opt. Soc. Am.* **59**, 1163–1171 (1969).

Prospects for life in Lake Nyos?

SIR—The active interest of the international scientific community in the disaster in the vicinity of Lake Nyos in Cameroon on 21 August 1986 is very encouraging. It is important to know the cause and mechanisms of this event so that it can be determined if others are likely to occur at Lake Nyos or other lakes of volcanic origin. I have, however, been struck by the lack of mention in the reports that I have seen in *Nature* and other science journals of the animal life in the lake.

For example, I am aware that the volcanic lakes of western Cameroon are well known to biologists as a natural evolutionary laboratory for fishes, mostly of the

family Cichlidae. As most of the 40 or so lakes have been isolated from each other for millions of years, new species of fish have evolved which are endemic to certain lakes. I have seen reports that Lake Barombi Mbo (or Lake Kumba) contains 11 endemic cichlids¹⁻³ and one endemic *Clarias*¹, Lake Barombi Kotto and Lake Mboandong five cichlids including two endemic forms (one species and one subspecies)³⁻⁵, Lake Ejagham two (possibly three) endemic cichlid species (ref. 6 and E. Trewavas, personal communication) and Lake Bemini one endemic cichlid⁷.

The important question is: what has happened to any fish that were living in Lake Nyos? If they did not survive the event, we would be comforted by the information that events of this magnitude occur very rarely in the western Cameroon volcanic lakes and have not happened at Lake Nyos for perhaps millions of years. On the other hand, if Lake Nyos is atypical and had no cichlids before the disaster it might be a sign that such events had occurred before and killed the fish. It would then be worth investigating if similar events in the past, particularly those associated with "legends of the Lake Monoun area, [where] there may have been at least three earlier cases of exploding lakes or mass death"⁸. There are reports that Lake Monoun had a population of fishes that were killed by the 1984 event, but it is believed that the fishes had never been recorded scientifically (E. Trewavas, personal communication).

If the fish population of Lake Nyos has not been seriously affected by the event, the tolerance levels of the fish for various physical and chemical impacts might be helpful in deciding which of the various scenarios postulated for the event could in fact have occurred.

A similar situation might exist for the distributions of other vertebrate and invertebrate species, which could provide additional data of importance to the examination of previous events and the prediction of future ones. I would be pleased to know if anyone has followed or is following this line of investigation and, if so, what results have been obtained.

WALTER SHEARER

The United Nations University,
Toho Seimei Building,
15-1 Shibuya 2-chome,
Shibuya-ku, Tokyo 150, Japan

1. Trewavas, E., Green, J. & Corbet, S.A. *J. Zool., Lond.* **167**, 41-95 (1972).
2. Trewavas, E. *Tilapia Fishes of the Genera Sarotherodon, Oreochromis and Danakilia* (British Museum, London, 1983).
3. Trewavas, E. *Bonn. zool. Beir.*, 146-192 (1962).
4. Fryer, G. & Iles, T.D. *The Cichlid Fishes of the Great Lakes of Africa: Their Biology and Evolution* (Oliver and Boyd, Edinburgh, 1972).
5. Trewavas, E. *Bull. Brit. Mus. Nat. Hist. (Zool.)* **26**, 329-419 (1974).
6. Thys van den Audenaerde, D.F.E. *Rev. Zool. Bot. Afr.* **76**, 349-356 (1967).
7. Thys van den Audenaerde, D.F.E. *Rev. Zool. Bot. Afr.* **85**, 93-98 (1972).
8. Kerr, R.A. *Science* **235**, 528-529 (1987).

Schrödinger's grand theme shortchanged

SIR—Perutz's¹ critical review of Schrödinger's *What is Life?*² focused only on the leaves and not only missed the forest but the trees as well. His review sifted through reductionist references of the past and dismissed the important themes developed by Schrödinger, who noted that living systems displayed two fundamental processes. One he called "order from order", and the other "order from disorder". Schrödinger synthesized and characterized the existing physical knowledge of the soon to be discovered DNA with the former and then proceeded to integrate the laws of thermodynamics with biology with the latter.

Perutz is correct in recognizing the influence of the earlier Timoféeff, Zimmer, and Delbruck paper³ on Schrödinger's determination of the possible size and stability of genetic material. But Schrödinger, as a physicist, should not be faulted for failing to have followed all the obscure literature in genetic and molecular biology being written at the time. His view of the genome, as an aperiodic crystal with unusual stability and coding capabilities was demonstrated later by Watson and Crick's analysis of ribonucleic acids and provided a framework that has led to some of the most important findings in biology today.

Schrödinger's most important and least studied observation was his "order from disorder" premise which links all biological systems with the expanded fundamental theorems of thermodynamics. He noted that, at first glance, living systems seem to defy the second law of thermodynamics. The law insists that, within closed systems, entropy should be maximized and disorder should reign at equilibrium. Living systems, however, are the antithesis of disorder as they display marvellous levels of order made from disorder. Schrödinger solved this dilemma by turning to 'non-equilibrium thermodynamics', recognizing that living systems live in a world of energy flux and that an organism stays alive in its highly organized state by taking energy from outside itself, or from a larger encompassing system, and processes it to produce a lower entropy state within itself. Life is a far-from-equilibrium dissipative structure that maintains its local level of organization, or negentropy, at the expense of the large global entropy budget.

Prigogine⁴ and Morowitz⁵ have extended thermodynamics into the realm predicted by Schrödinger, and new advances are rapidly being made in biology that are firmly rooted in this new thermodynamic paradigm⁶. All living systems appear as quasi-stable, far-from-equilibrium dissipative structures. These systems evolve via bifurcations and autocatalytic non-

linear processes, and the development of entropy production, growth, complexity and cycling and thermodynamically constrained biology may soon add new dimensions to physics (and thermodynamics in particular), as Schrödinger suggested it might.

Schrödinger was influenced by the Boltzman school and both men's work was based on statistical interpretations of Nature. Schrödinger's genius lies in the way he was able to describe the important microscopic features of life within the confines of quantum mechanics and statistical thermodynamics. Perutz contends that Schrödinger's essay was not original work, and its chief merit was the popularization of the Timoféeff paper. I am sorry that Perutz missed Schrödinger's grand theme of the possible description of biologic processes within the framework of modern physics.

ERIC D. SCHNEIDER

Chesapeake Biological Laboratory,
University of Maryland,
Solomons, Maryland 20688, USA

1. Perutz, P.M. *Nature* **326**, 555-558 (1987).
2. Schrödinger, E. *What is Life?* (Cambridge University Press, London, 1946).
3. Timoféeff-Ressovsky, N.W., Zimmer, K.G. & Delbruck, M. *Nachrichten aus der Biologie der Gesellschaft der Wissenschaften Göttingen* **1**, 189-245 (1935).
4. Prigogine, I. *Introduction to Thermodynamics of Irreversible Processes* (Wiley, New York, 1955).
5. Morowitz, H. *Energy Flow in Biology* (Academic, New York, 1968).
6. Ulanowicz, R. *Growth and Development: Ecosystems Phenomenology* (Springer, New York, 1986).

Chain flexibility and antigenicity

SIR—I bring your attention to a misconception about antigen-antibody interactions that continues to be propagated in the literature, most recently in an article in *Nature* (326, 358-363; 1987). There has never been evidence that "the antigenic regions of proteins are located in the more flexible chain segments". Rather, from limited data, it was suggested that continuous epitopes correspond to flexible regions (*Nature* **311**, 123-126; 1984). Even if this interpretation is correct, it is not known whether discontinuous epitopes, which probably make up the bulk of the antigenic regions on a protein, would also be found in the flexible regions. Moreover, further analysis of the antigenic sites on the tobacco mosaic virus protein (the key protein examined in correlating flexibility with continuous epitopes) has shown that the entire polypeptide chain is antigenic (*EMBO J.* **4**, 1231-1235; 1985). It is, therefore, unlikely that polypeptide chain flexibility plays a major role in antigenicity if it has anything at all to do with it.

RONALD JEMMERSON

Department of Microbiology,
University of Minnesota,
Minneapolis,
Minnesota 55455, USA

The men behind the bomb

Lawrence Badash

The Making of the Atomic Bomb. By Richard Rhodes.
Simon & Schuster: 1987. Pp.886. \$22.95, £18.

To Niels Bohr, complementarity was not a term to explain only the dualism found in nature; it applied equally well to the impact of nuclear weapons. The atomic bomb would destroy, but it would also end war, for it would turn war into national suicide.

Bohr had reached this insight at least by 1943, and he acted upon it by attempting (unsuccessfully) to persuade Roosevelt and Churchill to inform the Soviet Union about the bomb project and thereby reduce the danger of a post-war arms race. Leo Szilard was the other visionary of nuclear energy. More than any other scientist, he had been responsible for the concept and creation of the nuclear weapons programme, and like Bohr he had an early awareness of the limits on their use. Szilard also fought (unsuccessfully) for scientists to have a voice in the political decision-making process concerning these bombs. For their efforts, both were considered subversive; Churchill wanted to have Bohr interned for the war's duration, while Leslie Groves, commander of the Manhattan Engineer District, wished the same for Szilard, and did actually keep him under surveillance.

The role of science in fashioning nuclear weapons and the relationship of science to government are the themes of Richard Rhodes's impressive book. From this framework are hung the details of scientific developments and, even more, the personal activities of the scientists. The story is taken not just from the discovery of fission but from early in the century, and the first 200 pages are a survey of the history of pre-war physics. The cast of characters is large and most are introduced by a substantial biographical sketch: Szilard, Bohr, Ernest Rutherford, Otto Hahn, Edward Teller, Werner Heisenberg, J. Robert Oppenheimer, Francis W. Aston, Ernest Lawrence, James Chadwick, Albert Einstein and many more. This emphasis upon personality makes the book highly readable, but the normally engrossing prose is somewhat strained at times, as in the formula-like introduction of new individuals and, more so, in the attempt to define their self-doubts, anxieties, hidden suffering or psychological traumas. If such troubles can be shown to have influenced science, as in Enrico Fermi's immersion in physics to block out his grief upon his brother's

death, then they are worth reporting. But the author rarely develops his case, and 'human interest' material of questionable accuracy becomes psycho-babble.

Rhodes devotes quite a lot of space to background material. This is generally well done, and includes a historical discussion of radioactivity, the nuclear atom, atomic numbers, isotopes, mass spectros-



Enrico Fermi (left) and Niels Bohr in discussion during a walk along the Appian Way in 1931.

copy, relativity, quantum mechanics, the neutron, accelerators and so on, always with the stress on the personal behaviour of the scientists concerned. One may quibble over some facts, for example whether Rutherford in 1907 thought that the atom was "substantially empty space"; over some colourful language, as in Rutherford roaring up from the Antipodes in 1895 to study in Cambridge; and over some interpretations, as in Rutherford's making his genteel poverty a virtue by refusing to seek wealth, and this anti-commercial attitude thereby leading to his famous 1933 comment that extraction of usable amounts of nuclear energy was "moonshine". Regarding wealth (through patenting an invention or becoming a businessman) and "moonshine", Rutherford did not have the entrepreneurial inclination, the experience or the connections to compete with Marconi in developing his wireless device commercially, and "moonshine" was always correctly qualified by "with the means at our disposal". Einstein and Robert Millikan often said the same thing, and it is creating a straw man to fault Rutherford for failing to predict a phenomenon of nature, fission, that was not actually discovered until over

a year after his death.

These, however, are minor matters, for overall the book is accurate and the characters are well drawn. Rhodes is a skilled author, and he weaves into his accounts of the scientific developments highly interesting discussions of the methods and philosophy of science, anti-Semitism, Chaim Weizmann and Zionism, the brilliance of Budapest and the march towards war of Nazi Germany. He continues with a clear description of the discovery of nuclear fission, the scientific reaction to it, efforts to inform the US government of the potential of a chain reaction, and the bomb projects in the United States, Britain, Germany, the Soviet Union and Japan. This story has been told often, though usually in fragmented form.

The great merit of Rhodes's book is that it embraces the large picture in substantial detail, though necessarily with less than a monograph for each part.

There is further merit in the points Rhodes chooses to emphasize and the questions he raises. Did the German scientist Walther Bothe, a man abused by the Nazis, deliberately calculate graphite's neutron absorption cross-section to be too large, thereby condemning their reactor to use heavy water which was virtually unobtainable? Because the American project was based so much on fear that Hitler might get the bomb first, and because Vannevar Bush, the science czar of the time, had raised the question of espionage

with the President in 1941, why was the United States so slow (late 1943) in beginning to gather information on Germany's accomplishments? Unlike their counterparts in Britain, scientists in America were quickly and rather consciously excluded from policy matters; was this Bush's grasping for power and Groves's paranoia over secrecy, or were (and are) there institutional flaws in the running of the United States? Rhodes calls the project "a separate, secret state with a separate sovereignty".

More fundamentally, the author highlights the ill-fated encounter in 1945 between Szilard and James Byrnes. Byrnes, who was soon to be appointed Secretary of State, shook the physicist by focusing on the bomb's value in making the Soviet Union more manageable. Szilard looked further ahead, to the time when the Soviet Union would also have such armaments, and sought policies that would minimize the danger then. Take Stalin into our confidence now, he urged; even better, do not use the bomb. Byrnes, ever the politician, pointed out that Congress would be aroused over spending two billion dollars on a project that yielded no known product. Further, without proof that the

S. Goudsmit/W. H. Freeman

money had been well spent, how could Congress fund post-war research into nuclear energy?

As we know, the products of Oak Ridge, Hanford and Los Alamos were dropped on Japan in August 1945. Rhodes not only tells of the missions, but he presents about 18 numbing pages, mostly of eye-witness accounts, of the destruction of Hiroshima and Nagasaki. Except for an epilogue that sketches the development of the hydrogen bomb, the story ends here. The book is less scholarly than the magisterial *The New World, 1939/1946*, by Richard Hewlett and Oscar Anderson, yet more accurate than Robert Jungk's well-written but flawed *Brighter Than a Thousand Suns*. Altogether Rhodes has

produced the most readable, exciting and just book to date that covers both the bomb and the preceding four decades.

The book is not polemical, but Rhodes does gently advance a message. The nation-state of modern times, he says, has become a "death organization". Science, the other major force to arise in the past few centuries, is the only structure strong enough to challenge the nation-state, for the bomb has made it vulnerable. Our visionaries, Bohr and Szilard, appear to have been more practical than the so-called hard-headed politicians. □

Lawrence Badash is Professor of History of Science in the Department of History, University of California, Santa Barbara, California 93106, USA.

Clever computers

Bernard Meltzer

Artificial Intelligence and Natural Man, 2nd Edn. By Margaret A. Boden. *Basic Books, New York/MIT Press, London: 1987. Pp. 576. Hbk £25.95; pbk \$14.95, £10.25.*

THIS is the second edition of a book which, first published ten years ago, was one of the best non-technical introductions to artificial intelligence. It deals also with psychological, philosophical and social issues related to the subject, and differs from the original only in having an added postscript chapter dealing with some of the developments of the intervening period such as logic programming, basic robotics, naive physics, non-monotonic logic, creativity and parallel distributed processing (PDP).

A putative third edition — in ten years' time, say — would probably have to be much more radically modified because, in Kuhn's terminology, a paradigm shift in the subject seems to have started already, with the promise of fast-gathering momentum. It is too early to give a precise definition of this shift, but a good starting point for indicating its character is to consider the title of a section of the author's first chapter, namely, "Computers Manipulate Symbols". Here Professor Boden's intention was to demolish a common illusion of the time, that computers can process only numbers, but the assertion in the title as such is possibly misleading and bears some examination.

A Martian, examining — let us say — the behaviour and inner workings of a robot actuated and controlled by an internal "symbol-manipulating" computer, might very well fail to discover any symbols whatsoever in the system, but only certain patterns of movement, sounds, connections, electric currents and so on, just as we would do in the case of a dog or a

lion. In other words, the symbols are not inherent in the computer, they arise only from our interpretation of certain patterns of physical processes. But even to say this is rather misleading. For it is not the case simply that we observe such patterns and then interpret them as symbols, but that the particular form the physical processes in the computer take is determined by design (of the hardware and software of the system), and such design might be done (and in the past has mainly been so done) using the high-level notion of symbols and symbol structures.

Today, however, impressive computer models of mental processes have been designed which make much less use of the concept of symbols. For instance, at a workshop in Geneva last year Sejnowski described a model implemented on a PDP computer, which from the reading of children's stories learned rules of English pronunciation (as well as their exceptions!). Models of this kind work mainly by interactive activations and inhibitions of a very large number of connected, simple units; so far from symbols being manipulated in the system, it appears to have been a positively mind-cracking job for Sejnowski, from observation of the activity of the model, to come up with any sort of detailed account of the processes going on in high-level terms.

The term "subsymbolic computation" has been coined for this kind of processing, though it is clear from the discussion above that in a sense all digital electronic computation is "subsymbolic". The distinction has to do only with the character of the design of the system: for instance, a learning program which resulted in the addition of a new item to the knowledge base in a serial computer would be predominantly of type "symbolic", while one which varied a set of connection strengths in a PDP computer would be predominantly of type "subsymbolic". It is probable that the functional design of our own brain and nervous systems is subsymbolic in this sense.

It might appear that the new paradigm is entirely the result of the advent of PDP computers. Certainly, the effort to exploit these massively parallel systems with their promise of very much greater speed, robustness and flexibility for cognitive modelling, compared with the von Neumann (serial) computer, has been determinative in this development.

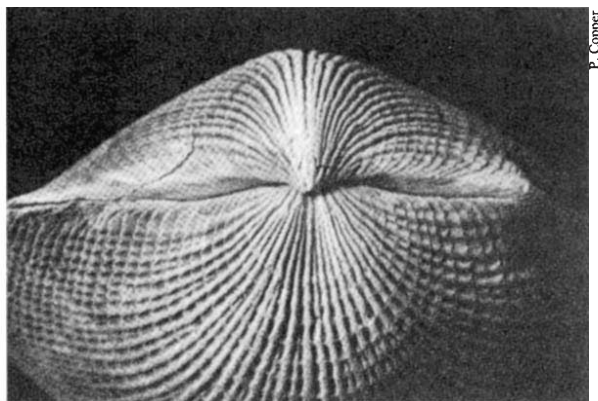
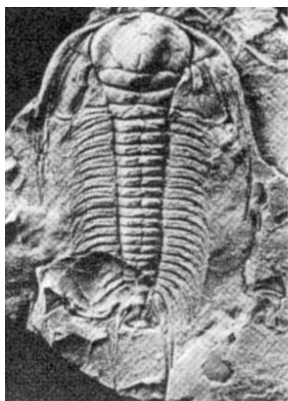
But there have been other, less palpable, factors involved. With its emphasis on symbolic representation, the von Neumann computer encouraged certain restrictive attitudes in the development of cognitive modelling. For example, from its beginnings at the end of the Second World War, artificial intelligence was dominated by the issue of the representation of knowledge, which — when one comes to think of it — is rather peculiar: an encyclopaedia is a good (symbolic) representation of knowledge, but hardly an intelligent system. The central issue surely is the process of thinking. Knowledge on the one hand is the result of such processes, and on the other hand is able to supplement sensory and motor experience in providing the raw material for them.

As for the symbolic representations themselves, they were in the main those suggested by natural, mathematical and logical languages: these are dominantly what Sloman termed "Fregean", in that predicates and functions are represented by the application of function symbols to argument symbols. However a map is also a symbolic representation, but one in which relations are represented by other relations, and it is only comparatively recently that the power of such non-linguistic "analogical" representations has been demonstrated on serial computers by cognitive models of naive mechanics and physics. PDP-type systems would implement such models much more effectively.

Finally, the "actor" model of computation proposed many years ago, in which computation is viewed not as the action of a central processor, but as the exchange of messages between data structures, was an anticipation of the new paradigm.

We see then that artificial intelligence, the computational modelling of cognitive processes, is still very much an exploratory and experimental discipline. This is the reason why it remains both exciting and, on occasion, frustrating — one can seldom can fall back on well-established principles and laws to guide research, as in some of the older sciences. However, Boden's book is, for the layman, still one of the most readable and certainly the most scholarly initiations to the subject. □

Bernard Meltzer, Via G. Galilei 5, Ispra (Va), Italy, was formerly Head of the Department of Artificial Intelligence at the University of Edinburgh, UK, and is now a consultant to the European Commission on basic research in artificial intelligence and signal processing.



A trilobite, *Paradoxides davidis*, of the Cambrian (left) and a brachiopod, *Gotatrypa orbicularis*, of the Silurian (right). Photographs reproduced from the newly published second edition of *Evolutionary Biology* by Douglas J. Futuyma. The publishers are Sinauer Associates, Inc., Sunderland, Massachusetts, USA, price is \$32.50; £29.50.

Working the patch

C.J. Barnard

Foraging Theory. By David W. Stephens and John R. Krebs. *Princeton University Press*: 1987. Pp.247. Hbk \$40, £25.10; pbk \$14.50, £9.10.

OPTIMALITY modelling in evolutionary biology has not had an altogether smooth ride. This has had little to do with the actual merits or otherwise of optimization models, but a lot to do with the perception of what they set out to achieve. Interpreted naively (or perversely) the notion of optimality appears to be simplistic, implying an unlikely state of perfection in a world shaped by chance and natural selection. Why it implies nothing of the sort has been spelled out often enough. But entrenched misunderstanding dies hard, and one of the purposes of *Foraging Theory* is to take the high ground in the advocacy and defence of the optimality approach. This it does with considerable skill. Foraging behaviour has proved by far the most fertile ground for the development of optimization models, partly because the concepts and mathematics of economic decision theory translate easily into models of choice by predators, and partly because foraging behaviour is richly amenable to controlled investigation in the laboratory.

The book begins with the fundamental components of optimization models and the development of basic average rate maximizing models of prey selection and patch exploitation from Holling's disc equation. Although the discussion quickly becomes punctuated with mathematics, the crisp and lucid text always keeps the reader in touch with the arguments and the flow is helped greatly by boxes containing mathematical derivations and explanatory asides.

The main part of the book is concerned with elaboration of the basic models. In

their simple form, the models are effectively unconstrained, assuming complete knowledge of the environment, unrestricted travel, instant recognition of prey and so on. The power of the approach is that biologically interesting features such as physiological constraints, trade-offs between competing activities and learning rules can be incorporated into the models as relatively digestible mathematical functions which allow the generation of precise and often counter-intuitive predictions.

Even with these modifications, however, the models still have some unwelcome limitations. They assume, for instance, that the world is completely certain and predictable, and that the optimal policy for the forager does not change as a function of its previous decisions. The chapter on risk-sensitive foraging injects some stochasticity into the forager's world and looks at choices based on probability distributions rather than certainties. That on dynamic optimization allows decision variables to depend on contingencies such as body size, past experience or estimates of future success which may themselves depend on the outcome of past decisions. In three final chapters the authors discuss the emerging interest in mechanisms (rules of thumb) that animals use in solving foraging problems, review the criteria for testing foraging models, and rehearse the main arguments and counter-arguments about the usefulness and validity of the optimality approach.

Foraging Theory is a *tour de force* of theoretical biology, written with care and precision by authors who thoroughly understand their subject. If anyone still seriously thinks that optimization models are "...naive, vacuous, tautological or just plain wrong..." (p. 206) they should read this book. I doubt they will think so once they have. □

C.J. Barnard is a Lecturer in the Department of Zoology, University of Nottingham, Nottingham NG7 2RD. UK.

Supplying the facts

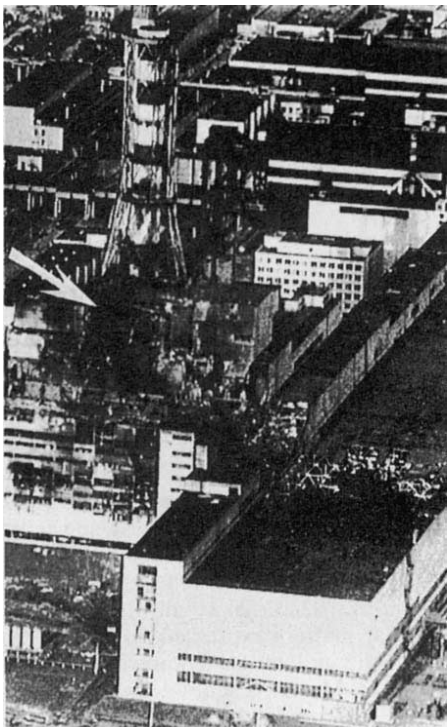
Murdoch S. Baxter

Environmental Radioactivity From Natural, Industrial, and Military Sources, 3rd Edn. By Merrill Eisenbud. *Academic*: 1987. Pp.475. \$55, £41.

MOST of the books reviewed in these pages are on specialized topics about which the average reader has little knowledge and hence no wish to pontificate. This text is a ready exception — every *Nature* reader, *everyone* even, has an opinion of some kind about nuclear reactors, nuclear weapons, nuclear discharges and their environmental and medical effects. Yet this also is a specialist field of science, needing detailed knowledge and understanding to support one's opinions. That, alas, most views on nuclear matters expressed nowadays are not validated by such knowledge is no fault of Merrill Eisenbud, who, for 24 years, has provided a textbook containing the basic facts and principles needed for sensible entry into the nuclear debate.

This, the third edition of *Environmental Radioactivity*, continues Eisenbud's contribution towards divorcing fiction from fact. As before the emphasis is on principles, the quality is high, the information concise, and the discussion powerful and balanced. Because Eisenbud is an authority of long-standing, he provides not only a current but also an historical view of the nuclear scene, and is well able to put events into their proper perspective. Happily, publication was sufficiently recent to allow inclusion, within the main text, of a full summary and discussion of the nuclear accident at Chernobyl in 1986, without which the book would clearly have been devalued.

There are some specialists who regard natural radioactivity as being of at least equal scientific and social importance to the artificial forms emanating from nuclear power plants, weapons and other technologies. Natural radioactivity is, after all, used in dating geological samples and archaeological artefacts; its heat output has helped drive evolutionary processes within the planet; its component nuclides provide tracers to study rates and mechanisms of natural processes; and its associated flux of ionizing radiation generally produces a radiation dose far in excess of that from man-made inputs. Eisenbud largely ignores this academic dichotomy — this is not a book which gives full weight to the naturally occurring radionuclides. Rather it concentrates upon artificial radioactivity: nuclear reactors, nuclear fuel reprocessing, nuclear weapons, environmental dispersion of radioactive waste products, nuclear waste management, nuclear fallout (short-term and



A view of Chernobyl nuclear power station. The arrow shows the spot where the accident occurred.

global), nuclear accidents, biological effects of radiation, radiological standards, safety, and risks and the public perception of them. Where naturally occurring radionuclides are discussed, it is their anthropogenic or technological enhancements that are emphasized. Though the title may therefore be misleading to a purist, the contents are primarily those of major popular interest — and common factual deficiency.

This is a masterly work which will be a valuable reference source for both the specialist and the general reader. For those such as myself who already have the first two editions, this latest version is sufficiently different to be a worthwhile investment. Some previous chapters and sections have vanished (surveillance methods) or shortened (radiation standards), while updating of text and references plus additions, such as that on Chernobyl, provide an up-to-date technical summary.

Just one thing troubles me — the author uses only the old units of radioactivity and radiation dose. These are scientifically obsolete, long-gone and buried. In teaching, research and even the media they have been replaced by the new SI regime of becquerels, grays and sieverts. Despite Eisenbud's explanation, in the preface, for his preservation of the old units (curies, rads and rems), their continued use will confuse readers new to the subject and irritate an older generation that is just congratulating itself on having mastered the transition. □

Murdoch S. Baxter is Director of the Scottish Universities Research and Reactor Centre, East Kilbride, Glasgow G75 0QU, UK.

Mapping the global gardens

Stanwyn G. Shetler

Floristic Regions of the World. By Armen Takhtajan. Translated by Theodore J. Crovello under the editorship of Arthur Cronquist. University of California Press: 1986. Pp.522. \$60.00.

AN essential framework for the comparative study of floras is a rational classification of the world's floristic areas. Such a classification is only as good as the plant systematics on which it is based, however, and floristic geographers have often been taxonomists or phylogenists first. Not surprisingly, then, Armen Takhtajan, a long-time student of floristic relationships and classifications, is also a leading authority on the phylogeny of the vascular plants. His new book presents the latest of many efforts over the past 150 years or more to elaborate a system of classifying floristic areas — phytochoria — of various ranks on a global scale.

Takhtajan has previously published several versions of his system. The present work is the most detailed yet, and is a revised and greatly expanded translation of his earlier book of the same title, published in Russian in 1978. It contains many new elements, including a much enlarged North American section by Arthur Cronquist, and more than 150 pages of appendix, bibliography and index. The appendix presents, appropriately, the author's latest version of his own phylogenetic classification of the living vascular plant families and higher taxa. This inclusion of his taxonomic macrosystem broadens the value of the book for systematists generally, while documenting the family system used in his floristic classification.

The author follows the traditional practice of subdividing the world into a hierarchy of mutually exclusive kingdoms, regions, provinces, districts and intermediate categories, with the floristic "region" being the main currency of the system. Six kingdoms (Holarctic, Palaeotropical, Neotropical, Cape, Australian, Holantarctic), 35 regions and 153 provinces, as well as a number of subkingdoms and subregions, are recognized, defined and characterized at least briefly in terms primarily of degree of endemism at the species and higher levels. Takhtajan sticks mainly to the floristic criteria, believing that geobotanical criteria should play a secondary role in phytochorionomy, and his characterizations often include relatively lengthy enumerations of taxa. Cronquist, on the other hand, uses geobotanical criteria liberally in his North American sections, and is more expansive in his interpretations. Some basic outline

maps are included, but the book really needed many more of them.

Takhtajan has travelled widely, and not surprisingly those regions he is most familiar with are very well covered. Users will be especially grateful for the floristic information about Soviet Eurasia, drawn from the Russian botanical literature and reviewed by Soviet specialists. It is a pity, though, that a "Cronquist" was not engaged to deal with the Neotropics; the discussions of these regions are sketchy, and convey nothing of the richness and complexity of their floras.

Floristic geography is not an exact science, and the theoretical grounds for the hierarchical classification of floras are debatable. Not all botanists share the author's enthusiasm for chorionomy in the formal sense, or his seeming faith that a truly natural system of choria can be eventually elaborated. As Takhtajan points out, floristic relations are usually complex, and the reality of the concept of "natural" floristic areas, not to mention a hierarchy of such areas, is relative at best; because degree of endemism is so dependent on size of area, it is difficult to avoid circularity in defining them unless independent data on the evolution and history of the flora can be brought to bear.

The book contains many individual insights into the origin and history of the regional floras, as understood in the light of current knowledge of plate tectonics, though one might have wished for a general discussion of plate tectonics and floristic geography in the introduction. Whatever the author thinks of vicariance biogeography, he takes no apparent notice of its arguments or methods. In the best traditions, his classification is the product of keen intuitive analysis, synthesis of a vast literature and a lifetime of personal experience in the study of the phylogeny and geography of plants.

Such misgivings notwithstanding, this is a remarkable compendium. Takhtajan not only provides researchers with a wealth of ideas about origin and dispersal for further investigation, but, as he himself stresses, he also sets forth a global perspective for the conservation of plants and their genetic resources. The book's faults are minor in comparison to its strengths, and a large number of biologists and conservationists will want to have a copy of it close at hand. □

Stanwyn G. Shetler is a Curator of Botany and Acting Associate Director at the National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA.

● Newly published by Cambridge University Press is *River Plants of Western Europe* by S.M. Haslam, with contributions and illustrations by P.A. Wolseley. The book describes and discusses the vegetation of watercourses of the European Economic Community, with an emphasis on distributional, community and historical ecology. Price is £75.00, \$125.00.

Uniquely West German research grants

J. Boeckh

The West German system of Sonderforschungsbereiche—special programme grants—has been in operation for 20 years. The system is unique and has many advantages.

FUNDAMENTAL and applied research in West Germany is funded mainly from the budgets of the federal government, the states (*Bundesländer*) and industry (Table 1). More than 70% of the total is spent by industry, about 14% by the universities, and about 13% by other public institutions (Tables 2 and 3). Most industrial research is directed to technology and application. Governmental research policy has focussed large-scale funding for fundamental and applied research on selected fields by operating institutes such as those for nuclear physics, space research, plasma physics, cancer, and environmental research. Fundamental and much applied research over the whole spectrum of disciplines is carried out by smaller public institutions and by government-funded scientific organizations like the Max Planck Society but the largest proportion is carried out in the universities (basically funded by the states), which employ more than 60,000 academics.

The rising load of educational and administrative tasks absorb an increasing amount of the working capacity of academic and technical staff, which results in a considerable shortage of personnel for research. Rising salaries and costs have reduced the research budgets of the universities and other institutes to a fraction of the sum listed in Table 1, a minimum for existence, causing a widening gap between demand and supply in exactly the institutions responsible for most of the fundamental research in the country, and where young researchers are educated and work, often for the most innovative periods of their careers.

The shortfall is partly provided by outside funding of university research. In 1985, these sources provided 2.3 bio. DM compared with about 1.5 bio. DM for research and teaching expenses from the universities' budgets. The *Deutsche Forschungsgemeinschaft* (DFG; German Research Association) provided 40% of this sum; the rest came from industry, federal and state governments, foundations like the Volkswagen foundation and federal scholarships for graduate and postgraduate students.

The DFG is the major non-governmental agency for distributing federal and state research funds to individual researchers, groups or institutions. Support is granted after review and recommendation by independent experts. Generally, an individual researcher is supported for two years with funds for equipment, ex-

penses, travel and personnel, but not the applicant's salary. To initiate and promote a special field a *Schwerpunktprogramm* (SPP; Priority Programme) may be set up, supporting some 15–30 researchers from all over the country for up to 5 years, with special emphasis on cooperation and communication. Other schemes, listed in Table 4, include those for training and support of young scientists.

In addition to these conventional support schemes, a unique programme for funding research in particular institutions, the *Sonderforschungsbereiche* (SFB), was initiated in 1967. At that time, research in many disciplines, especially in technology, medicine and natural sciences, had become extremely expensive, and the investigation of major scientific problems was increasingly demanding cooperation between many specialists in a range of disciplines. In addition, the development of research outside the universities and the founding of many new universities caused a drain on the reservoir of talented and competent personnel. It became difficult to achieve a concentration of leading scientists in a particular discipline at a given university. Moreover, it was becoming apparent that research in Germany was falling behind international scientific development in several fields.

For all these reasons, it seemed necessary to concentrate resources and manpower for research to a limited number of fields in the most suitable places, to strengthen research in the universities and to recruit high-quality staff at all levels. These worries, common among politicians and scientists, were taken up by the *Wissenschaftsrat* (science council), an advisory committee of scientists and delegates from various federal and state offices appointed by the president of the Federal Republic to advise the government on scientific affairs. In 1967, this

committee recommended the formation of centres of excellence in research at universities where certain preconditions promised fruitful development. The scheme they proposed has, with some changes, been operating ever since.

The SFB scheme

An SFB, literally translated as 'special research area' or 'special collaborative programme' is a cooperative of scientists at a single university formed to undertake a large-scale investigation of a major problem. The university administration endorses the programme and gives it high priority. The term SFB is used for both the cooperative and the topic investigated.

An SFB is intended to be a nucleus for the development of a particular research field in a university. SFBs bring in additional resources and should be able to attract well-qualified staff and students who in turn contribute to the scientific standards of the university and the field. They also lead to a division of labour and the selective concentration of potential within particular universities. The effort and expenditure required to run SFBs limit their number, both in a university and throughout the country but, because they cover a wide range of subjects, SFBs make a considerable contribution to the profile and diversity of the scientific landscape in West Germany.

The major characteristics and special features of SFBs are:

- Long-term and large-scale funding on a solid foundation but with a limited lifetime of 9 to 12 years, rarely more.
- Funding of a local 'critical mass' of research units with sufficient expertise and quality.
- SFBs are funded from a budget provided by federal and state governments.
- Involvement and responsibility of the

Table 1 Research funding in West Germany in 1986 (in million DM)

Sources		Budgets	
Federal	13,225	Universities	7,785
States and other public	7,930	Other institutions (Table 2)	7,475
Industry	34,490	Industry	39,685
Foundations	205		
National	55,850	National	54,945
International	610	International	1,515
Total	56,460	Total	56,460
ca. US-\$	25,900		

The published budget figures are an estimated 30% of the total sum spent for universities (ca. 22 bio. DM) including their hospitals. Actual sums for teaching and research are only a fraction of the total. National funding for research is 2.9% of the gross national product.

Table 2 Federal, state, and other public funds spent for research by non-university institutes and by the DFG in 1985 (in million DM)

Large research institutes	2,660	
Smaller federal and state institutes	1,111	
Societies	1,238	
Libraries, museums	321	
Others	480	
Subtotal	3,150	3,150
DFG		965
Total	6,775	

university administration and the relevant government ministry.

- Regular external in-depth review of progress and requirements but at 3 yearly intervals instead of the more usual 2 years for standard grants.
- Intense internal cooperation and evaluation of progress; all participants have joint responsibility for the performance of the whole SFB.

All decisions about setting up, funding, evaluation (including review) and administration of SFBs is in the hands of a special branch of the DFG. The *Wissenschaftsrat*, state and federal offices and the *Westdeutsche Rektorenkonferenz* (the conference of heads of the universities) are involved in the decisions.

Establishing an SFB

The initiative for an SFB comes from a group of scientists at a university, who agree to cooperate on a common theme. In the early days, some of the research themes were defined very widely, for example "Molecular Biology of the Cell" or "Plasma Physics" but more recently topics have been much more circumscribed (see Table 4), with the consequence that there are now often fewer, more coherent projects in a single SFB.

There is no external policy or direction for the selection of themes, nor is there a preconceived distribution of SFBs by discipline or geographical location. Nevertheless, the distribution has been fairly constant.

The formulation of the theme and objectives sets the criteria for participation. All members of research units with appropriate interests in the university are potential members, which should ensure the participation of all useful units and enforces cooperation. A major policy of the scheme is to restrict membership of an SFB to members of the home university and the immediate neighbourhood, so that research institutions in the same metropolitan area may be included. As the object is to develop and support a local nucleus and to reward a university for its role and developmental policy, physical proximity of the participants is deemed essential. The inclusion of institutions outside the university facilitates cooperation between traditionally separate research

disciplines, for example, between applied and fundamental sciences in technology or medicine. It also helps to contribute to the critical mass essential for investigating the common theme and helps to recruit important expertise from neighbouring fields or auxiliary disciplines. Composition and internal structure of an SFB are carefully monitored by the reviewers and the DFG, and restrictions or extensions to the draft programme may be strongly recommended.

The selection of the final group of research projects and the members of the proposed SFB, the development of an internal structure and the preparation of a budget require constructive discussion. Often this process is seriously hindered by dispute, friction and other problems of group dynamics. Finally, the members adopt a proposal for establishing a programme for an initial three-year funding period. A statute regulating internal structure, self administration and control of progress is drawn up. The statements required for normal grant proposals, covering objectives, preliminary work, detailed research plans and justification of budgets are also required. The responsibility for and decisions about the whole SFB, down to the budgets of the individual projects, are shared by all the members. The university administration decides if it will adopt this proposal and it guarantees, within certain limits, to provide the basic establishment. After approval by the state ministry, the application is submitted to the DFG by the university. The involvement of and acknowledgement by ministry and university are intended to guarantee stability for the SFB, which is especially important during times of reduced public budgets or changing policy.

The proposal is subjected to an in-depth review. The DFG nominates a panel of about a dozen recognized scientists, each an expert in one of the fields of the proposed investigation. Although its members may change, this team accompanies the SFB for the whole of its life. The independence of the review team, and the care and intensity of the review procedure are key elements of the whole SFB programme, providing the basis for confidence in the system and justification for the considerable expenditure.

To evaluate the proposal the team visits the university, where it learns the details of research, local facilities and budgets through plenary presentations or personal meetings with staff. The whole proposal is discussed at a joint meeting attended by the panel, SFB staff and representatives of the university administration, ministry, the DFG- and the *Wissenschaftsrat*. Strengths and weaknesses are openly discussed, the budget verified and local conditions determined. During the whole review, the proposers have the opportunity to reply and explain to an extent that is unusual in the evaluation of research proposals. It is also unusual that administrators are exposed to actual research matters and the detailed concerns of research groups. They find themselves confronted with critical questions about local prerequisites and with recommendations for the improvements that are considered essential for setting up the SFB. Nor do scientists normally have their research evaluated in an open and detailed discussion in front of their colleagues and administrators. This elaborate and often stressful procedure ensures the review has maximum transparency and objectivity.

The final report of the review panel, which is worked out in closed session with no SFB members present, includes recommendations for establishing the SFB, and about its structure and funding, down to single items in the budgets. The subject and the objectives and perspectives are judged, and the quality of the whole SFB, of each research project and of the work of the scientific members is assessed. Once approved by the *Wissenschaftsrat*, the final decisions about setting up the programme and the first three-year period of support are made by a DFG executive committee composed of elected and delegated scientists, and representatives of federal and state governments, the *Wissenschaftsrat* and the rectors' conference. The decision includes a conditional recommendation, and basic changes and restrictions may be required. At three-yearly intervals, the review panel evaluates research reports and new proposals for continued support and, provided there has been satisfactory development, the programme is funded for up to four or five consecutive three-yearly periods.

Table 3 DFG budget in 1986

Programme	Number of grants	Million DM
Individual projects	4,646	475.4
SFBs	163	315.2
Projects in 117 SPPs	1,532	181.4
Libraries		21.6
Scholarships and Heisenberg research fellowships	248	25.7
Small research groups	24	17.7
International relations		18.9
Other		20.1
Total		1,076.0

After 12–15 years, the role of an SFB as a springboard should have been accomplished, with the field sufficiently well-established at the institution. The topic may also be somewhat exhausted and new initiatives lacking. This is the time for new fields to be developed and new constellations of scientists supported, although members of an old SFB may join initiatives for starting another programme.

Funding figures

An SFB programme finances many projects and provides its members with their major source of finance for research. Support includes funds for scientific and technical co-workers on temporary contracts, equipment, research expenses, travel and guests. Each project in the programme has its own approved budget but some equipment and funds, for example, for travel and guests are generally available. There is considerable internal flexibility, even to the extent of mutual cover between budgets for salaries, expenses and equipment that is unusual for fiscal budgets.

The budget is administered by the financial department of the university; all funds must be spent for SFB purposes with no deduction for overheads. There are on average four SFBs in a university, although two of the technical universities (Aachen, Munich) have 10 each. The basic staff of an SFB is above Ph.D level, with a maximum of 75. From 1968 to 1985, 230 SFBs have been established and 95 terminated, with a total expenditure of 3,317.8 million DM (Table 5). A maximum of 162 programmes was reached in 1985 when 29 began their first triennium and 27 were terminated. The total budget for that year was 303 million DM. The average budget for an SFB is about 2 million DM per annum, with a maximum around 5 million DM, with up to 75% spent on salaries. The total budget has increased over the years as the number of SFBs has grown but inflation has resulted in a loss of 25% in purchasing power between 1978 and 1984 (Table 6). The total SFB budget is a little less than 30% of the whole DFG expenditure and in 1968 was 4.1% of the total research budget for universities (including their hospitals). 75% of the SFB budget is covered by federal funds, 25% by the states.

Internal structure

As the members and projects of an SFB share risks and profits, a sound internal organization is vitally important. A statute adopted by the first members of the initiative defines and regulates membership, the overall theme, principal objectives and perspectives, internal control of progress, administration, initiation of new projects, forms of application for funding, and the sharing of responsibilities. Basic decisions are taken by the assembly of members, which consists in some SFBs solely of

project leaders, in others of all participating scientists at or above Ph.D-level (except those employed by the SFB). A 'speaker' represents the SFB to the world and oversees the management of the programme, sometimes with the assistance of a board and/or committees. The speaker and committee are elected by the members and report to the assembly.

Each SFB consists of several research units (projects or sub-projects), each headed by a project leader, generally a member of the university staff. Project leaders are SFB members and in the SFB organization are responsible for inner and outer control of the SFB. Except central equipment available to many units and money for travel and guests, individual projects apply directly to the SFB for funding.

Major features of internal organization are regular progress reports on the projects, quality control, for example, during discussion of the grant proposals, and cooperation between the projects.

SFBs form special entities within a university and modify the traditional hierarchy of scientists in German universities. The projects in an SFB are carried on in several departments and faculties, and members and project leaders vary in their university ranks. Within the SFB, however, all members and project leaders have equal status, sharing basic decisions and self-government. Funds and personnel are at the disposal of project leaders, who are not always at the professional level, but often include more junior associates, some in non-tenured positions.

SFB budgets provide for a substantial number of temporary pre- and post-doctoral jobs, limited to a 3 or 5 year maximum, depending on the level. These younger collaborators are particularly significant because research benefits greatly from the accumulated talents of young scientists, who are often in the most fertile and innovative period of their careers. This policy, however, is not without problems partly because of the poor chances for younger workers to establish permanent careers in fundamental research. This also causes difficulties with recruiting young talent into SFBs. Consequently, there are special conditions for the employment of temporary personnel. It is tempting to keep scientists once they are familiar with the approach and the methods of a project and have become able co-workers and welcome producers of valid results. But SFBs have a major responsibility to see that this group remains mobile and flexible, to challenge them to expose themselves to different kinds of problems and to help to establish them in independent careers.

The expertise that accumulates in an SFB through the interactions between the various disciplines is enriched by distinguished guest researchers and seminars

Table 4 Examples of disciplines in which SFBs were started in 1986

Coronary heart diseases: Prevention and therapy for acute complications.
Protein phosphorylation and intracellular control of membrane processes.
Processes relevant to climate in the ocean-atmosphere-cryosphere system.
Molecular microelectronics.
Prevention and intervention in childhood and youth.
Social history of the modern bourgeoisie.
Internationalization of economy.
Aesthetics, pragmatics and history of television media.

with international participants. The climate of intense discussion provides an optimal environment for high-quality training. Several SFBs have created special post-graduate and postdoctoral training programmes with workshop seminars, and training on the job. Especially in fundamental research fields, it is seen as important that younger collaborators should not be retained for too long, but they must be prepared, trained and motivated for careers in other fields where job prospects are better.

Success of the system

The SFB system is very expensive and yet supports only a fraction of the scientific community and a selection of research fields. Such selectivity must be justified by success to counterbalance apparent disadvantages, and by modification in response to criticism.

Critics of the system raised several objections:

- The procedures of starting and running an SFB involve a lot of bureaucracy as well as placing a burden of committee work and administration on SFB members. Not all of the large number of recognized experts recruited to the review panels for over 150 SFBs can be found outside the system, leading to the absurd situation where reviewers and reviewed exchange roles, which could lead to loss of independence. This problem also applies to the review of other funding programmes.
- Innovation and progress in science depend basically on individual researchers, who should be funded with specially tailored support, with maximal independence and flexibility for individual development. They should not be lured by large-scale support and the potentials for cooperation and protection into a network of themes and objectives that contain compromises, or into a group of mixed quality. The problems of group dynamics, including inherent inertia and friction, and the elaborate self-administration procedures absorb too much energy and research time, and reduce flexibility and degrees of freedom.

The integrated approach might lead opportunistic scientists to abandon or redirect their own research programme, as well as to the establishment of new hierarchies and dependencies between scientists. Because good projects have to carry others to achieve a critical mass, the quality of research may be evened out rather than raised.

- The SFB system leads to distortion of competition and might become subject to a planned economy. Because the reviewers and the decision-making committees are part of the establishment that dictates the direction of major efforts, new or unfashionable themes and fields might fail to be accepted. SFBs and their members have a privileged place in their institutions, leading to discrimination against other members of the institution and reducing their chances. Because SFBs are local groups, they are excluded from collaboration with groups working on similar topics elsewhere, which may retard the development of some fields. The careers of established workers may be enhanced at the expense of the valuable but non-conformistic individual.

- Because of these disadvantages and the high costs of the system, it might be more effective to subsume the SFB budget into the well-proven programmes for funding individuals or individual units, or for the supraregional support of particular research areas (Schwerpunkt programme).

Advocates of the system point out:

- The SFB system allows continuous and comprehensive investigation of major research problems with the development of large scale approaches. The cooperation, often interdisciplinary, between many research units and the accumulation of expertise and equipment provides a good basis for risky approaches and long-term methodological developments. The proximity of collaborating units that have a vital interest in mutual success as well as improved facilities and long-term funding give SFBs an advantage over

geographically widespread research programmes supported by conventional funding. Because of their restricted lifetimes and the necessity to regenerate by forming new SFBs, the system has a flexibility and adaptability superior to that of permanent research departments in universities and scientific societies.

- SFBs have become an essential part of university research. The research is generally at international standards, and several SFBs are among the leading international groups in their fields. Many important and contemporary problems continue to be investigated successfully and essential new approaches developed. Many SFB members have received prestigious appointments and younger collaborators definitely enjoy better chances in a very competitive job market.

- Both for research units and individual participants, the system clearly has benefits other than purely financial. Intense interaction during the preparation of proposals, in the course of research and during internal assessment all considerably improve concepts, strategies and the performance of projects.

- Because the limited increase in government support to the universities cannot compensate for inflation and increased costs, outside funding becomes more important. The total research budget of universities can be increased by several million DM a year through SFB funds and equipment purchased by SFBs becomes university property. The universities also profit because SFBs attract well-qualified staff and students, one reason for the overt competition between universities to establish SFBs. Several universities have moved away from equal distribution of resources and now provide extra funds for special areas and for groups which recruit outside support on the basis of external quality control, so special consideration should be given to SFB members.

- The SFB scheme is not only an important special purpose alternative in the Ger-

man funding system but also a factor in the research and job policies of universities, and so helps to mould the contours of the scientific scene of the country. In the traditional universities, SFBs have helped to overcome borderlines between departments and disciplines, and to form connections between universities and other scientific institutions, as well as between fundamental and applied research. At several of the universities founded after 1964, appropriate appointment policies meant that SFBs were easily set up, so that research was well established early in the development of these universities.

Conclusion

No overall evaluation of the success and impact of the SFB programme has yet been attempted, and it will be difficult to achieve such an assessment. Now, 20 years after the start of the programme, most of the earlier-established SFBs have already terminated. A lot of experience must have been gained by the Deutsche Forschungsgemeinschaft, foreign scientists, editors of recognized journals, scientific societies and, last but certainly not least, by the members of the review panels. In its recently published report on the first 15 years of the programme (*Sonderforschungsbereiche 1969-1984*, eds Stackmann, K. & Streiter, A., VCH, Weinheim; 1985), the Deutsche Forschungsgemeinschaft itself proposes such an evaluation as necessary and part of the justification of the system.

From the author's experience of several funding schemes run both by the DFG and other bodies, he agrees with many colleagues from both inside and outside the SFB system that the advantages of the programme by far outweigh the disadvantages. Most of the promises listed above have to a large extent been fulfilled. Some of the disadvantages are inherent and compromises have to be made to optimize the gain.

The SFBs have an impressive and internationally respected record of research and their success in involving young researchers in high quality research and in promoting their development is well recognized. They have become the focus of research in certain fields and nuclei of recognized research in many universities, where they attract good researchers. Thus, the SFB system provides a substantial component of research at West German universities, and is acknowledged as an important complement to more routine sources of research funding. □

Table 5 Development of the SFB system: numbers of SFBs in different disciplines for the years 1968, 1975 and 1985

	1968	Year 1975	1985
Humanities (education, history, linguistics, literature, psychology, political sciences, social sciences)	4	19	19
Biomedical sciences (agriculture, biology, medicine, veterinary medicine)	8	45	65
Natural sciences (mathematics, chemistry, geosciences, physics)	3	23	39
Technical sciences including architecture	3	29	39
Total	18	116	162

Information for data presented in Tables 1-5 and in the text is from Annual Reports (*Tätigkeitsberichte*) of the Deutsche Forschungsgemeinschaft, D-5300 Bonn 3; Sonderforschungsbereiche: Grundlagen des Förderungsprogramms und Verfahrensregeln, Deutsche Forschungsgemeinschaft (1982); Stellungnahme zur Entwicklung des Programms der Sonderforschungsbereiche, Wissenschaftsrat Drs 6774/85, Berlin (1985); Drittmittel der Hochschulen, Wissenschaftsrat Drs 7423/86, Köln (1986); Eckdaten zur Lage der Hochschulen, Wissenschaftsrat Drs 7674/87 Köln (1987); Bundesminister für Forschung und Technologie, Faktenbericht 1986 zum Bundesbericht Forschung, Bonn, Mai 1987; Battelle Institut Frankfurt Forschungsbudget, 1987. In 1968 1 US\$ = 4.0 DM; in 1975 2.4 DM, in 1985 3.4 DM and in 1987 1.8 DM.

J. Boeckh is a professor in the Institut für Zoologie, University of Regensburg, FRG and is speaker of SFB4 in the University.

Evolving radio structure of the binary star SS433 at a resolution of 15 marc s

R. C. Vermeulen*, R. T. Schilizzi†, V. Icke*, I. Fejes‡§ & R. E. Spencer‡

* Sterrewacht Leiden, Postbus 9513, NL-2300 RA, Leiden, The Netherlands

† Netherlands Foundation for Radio Astronomy, Postbus 2, NL-7990 AA, Dwingeloo, The Netherlands

‡ Nuffield Radio Astronomy Laboratory, Jodrell Bank, Macclesfield, Cheshire SK119 DL, UK

A comprehensive series of VLBI (very-long-baseline interferometry) observations of SS433, carried out over an 11-day period on the European VLBI Network at 5 GHz, reveals the motion and evolution of a series of faint 'blobs' ejected from the system. These blobs are observed to brighten at a substantial distance from the binary star.

THE object SS433 features moving emission lines in its optical spectrum, which show redshift and blueshift simultaneously. This behaviour has been interpreted¹ as being due to the ejection of gas with a speed of $0.26c$ in two antiparallel directions. According to the generally adopted kinematic model, as described in the review by Margon², the ejection axis precesses with a 162.5-day period³. The object is generally thought to be a binary system, containing a massive early-type star losing mass to a thick accretion disk which surrounds a compact object². This object and the disk surrounding it are likely to be responsible for the acceleration and collimation of the ejected gas (ref. 4).

Radio observations have confirmed the existence of corkscrew-shaped trails. All evidence to date is consistent with the notion that the gas moves ballistically; thus, the corkscrews are not the path traced by the flow, but rather the instantaneous locus of gas clouds that, individually, move in straight lines away from the central source. Very Large Array (VLA) measurements⁵⁻⁸ on arcsecond scales showed that the precession cone half angle is $\sim 20^\circ$, and that the angle between the precession cone axis and the line of sight is $\sim 80^\circ$. The radio emission on this scale sometimes resembles an extragalactic double-lobed source, which shows knots corresponding to projection cusps in the corkscrew locus, on either side of the unresolved core⁹. Fits to the observed ejection pattern yielded a distance estimate of 5.5 ± 0.5 kpc⁹.

Knots are also clearly in evidence on the 100 marc s scales accessible by MERLIN¹⁰. Fits to the ejection pattern on these scales yielded a distance estimate of 4.9 ± 0.2 kpc. This work always showed a symmetrical structure, in which knots form, in pairs, at the time of radio outbursts. The knots move at constant velocity, decaying in brightness according to a power law with a time constant of 35 days.

Radio structure on a scale of 10–50 marc s was discovered in 1979¹¹. Since then, a substantial number of VLBI observations have been made. Early work showed that at 2,990 MHz there is a 30 mJy unresolved (~ 2 marc s) core¹²; at 10.65 GHz, a lower limit of 5 marc s to the size of the core component at the flux density limit of 120 mJy was found¹³. In addition, two groups¹²⁻²¹ have studied the extended emission on the scale of tens of milliarcseconds. Their efforts have shown that there are almost always well-collimated, blob-like structures, which usually trace the locus predicted by the kinematic model². The positions and proper motions of the knots yield the same values of the gas speed and of the distance to SS433 as those found from VLA and MERLIN maps.

We present the results of a further set of six closely spaced observations, but with a much improved sensitivity and baseline coverage. A set of optical spectra obtained in parallel will be reported elsewhere.

Observations and data reduction

In May 1985, we observed SS433 six times at intervals of 2 days, using the European VLBI Network (EVN) at 4,990 MHz. The Mark III recording system²² was used in Mode E: we observed a 14-MHz band in left-circular polarization. Table 1 presents details of the telescopes available. A rubidium oscillator frequency standard was used in Jodrell Bank, and the other stations used a hydrogen maser. The source was not detected on all available baselines at all times due to a number of factors which conspired to limit the available array. The 6-cm radio flux density of SS433 was unusually low: 0.3 Jy. Here, we will refer to each particular map by its sequence number, as defined in Table 1.

All data were correlated at the Mark III processor operated by the Max Planck Institute in Bonn, FRG. The coherent integration time was limited to 2 minutes, in view of phase noise losses due to the use of a rubidium oscillator in Jodrell Bank, a poor phase calibration signal recorded in Westerbork, and the non-recording of this signal in Medicina.

Following the procedure described by Cohen *et al.*²³, we applied an *a priori* calibration to our data based on telescope gain curves as a function of elevation, and on system noise temperatures which were typically measured every half hour. As in previous observations²¹, we observed the calibrator source PKS2121+053. By requiring internal consistency in the calibrator dataset, we could determine time-dependent corrections per telescope. We were able properly to recalibrate our SS433 datasets for source risings, but we failed to find consistent factors at setting. The affected datapoints were all eliminated before proceeding.

We also carefully inspected our calibrated datasets by eye for any irregularities. Bad points were deleted. For questionable data sections, we have run tests in which we compared hybrid maps made with and without such data, and the way these maps reproduced the observed visibility data. Hence we are satisfied that none of the features we will comment on this paper are present simply because we either retained or removed a particular data section.

For all six observations, we have used, independently, the NRAO AIPS package (mapping program MX), the hybrid mapping programs in the Caltech VLBI Software Package (courtesy of T. J. Pearson), and the Caltech Maximum Entropy program (due to Skilling and Gull²⁴). The features in the maps which we discuss in the following section emerged with all three methods. We base our further analysis on the AIPS maps, which were not only superior in reproducing the visibility data, but also required the least amount of windowing to suppress spurious features due to sidelobes.

Our hybrid mapping procedure was always started by assuming an unresolved point source for the first self-calibration. All important features emerged immediately in the first iteration.

§ Permanent address: Satellite Geodetic Observatory (Penc), Hungary.

In the first three hybrid mapping cycles, the major features found were used only for phase self-calibration. We then applied a fourth and fifth cycle, in which the sensitivity of each station was allowed to be changed by a constant factor. Typically, adjustments in sensitivity of up to 4% were made in cycle four, with no further corrections in cycle five. We had to apply a window that cut off the major sidelobes located further than 50 marc s to the north and south of the main structure, well out of the way of the precession trail. The dynamic range of the final maps is typically 50–1, and was limited by the sparseness of the UV-plane coverage; features below this level anywhere in the map could be spurious.

The resultant maps are shown in contour form in Fig. 1a–f. Equal length scales have been used for each map, for display purposes only. The restoring beam shown in Fig. 1 is the ellipse which best fits the half-power contour of the dirty beam. Map 6 is based on a very small time interval for which a triangle of baselines was available. The resultant map should therefore be treated with great caution, even though the structure shown did in fact emerge consistently from all three mapping methods used.

Figure 1 also shows the projections on the sky of the beam trajectories as predicted by the five-parameter kinematic model², when adopting a distance of 5 kpc. At the time of our observations, the predicted ejection angle was almost at right angles to the line of sight. Figure 2a–f shows crosscuts through the images, along the curved locus. To reduce the pixel-to-pixel noise, the flux densities shown represent averages over a 15 marc s wide strip (~ 1 beam), placed symmetrically around the kinematic model curve. Since we lack the absolute astrometric positions of our images, we have centred the model origin on the bright, unresolved core in our maps, which stands out clearly in Fig. 2. Open circles are drawn along the kinematic model curve at two-day intervals (equal to the time between the observations).

Daily measurements of the integrated radio flux density at 408, 843, 2,700, 5,000 and 8,100 MHz showed that SS433 was in a generally quiescent state which was interrupted only by a small flare (50 mJy at 5 GHz) on JD 2446201. The spectral index was always -0.65 ± 0.1 , in agreement with values found by previous investigators⁹ (see also R. Fielder *et al.*, manuscript in preparation). Exosat flux density measurements at JD 2446202 and JD 2446212 also revealed no flaring.

Critical examination of the maps

After the careful calibration and mapping procedure outlined above, a number of characteristic features clearly emerged from maps 1–5.

The maps contain a strong and variable unresolved (< 10 marc s) core, which we take to coincide with the stellar source in all maps. We list its flux density in Table 2. In maps 3, 4 and 5, the core structure clearly has extensions along the predicted precession locus²; these may correspond to subsequent unresolved ejection episodes. We are unable to distinguish a blob convolved with the synthesised beam from a continuous flow in this case.

In addition, the maps show a series of discrete emission patches: Fig. 2 shows a high brightness contrast between the knots and the regions in between. The maps are in general quite symmetric, though the detailed flux density distribution east and west differs. This supports our choice of the brightest, central knot as the core. Map 2, where there is more flux density in the eastern than in the western beam, is an exception. Note that the emission is unresolved perpendicular to the trace, whereas along it, the features are much broader. In map 1, we see knotty emission immediately adjacent to the core. By map 3, this emission has clearly separated from the centre; at the 50–1 dynamic range of our maps, there is no emission at a distance of ~ 30 marc s on either side of the central feature. None of our maps displays any emission beyond ~ 100 marc s from the centre.

When the brightest points in Fig. 2 are connected from frame to frame, we get lines that diverge at (at most) three quarters

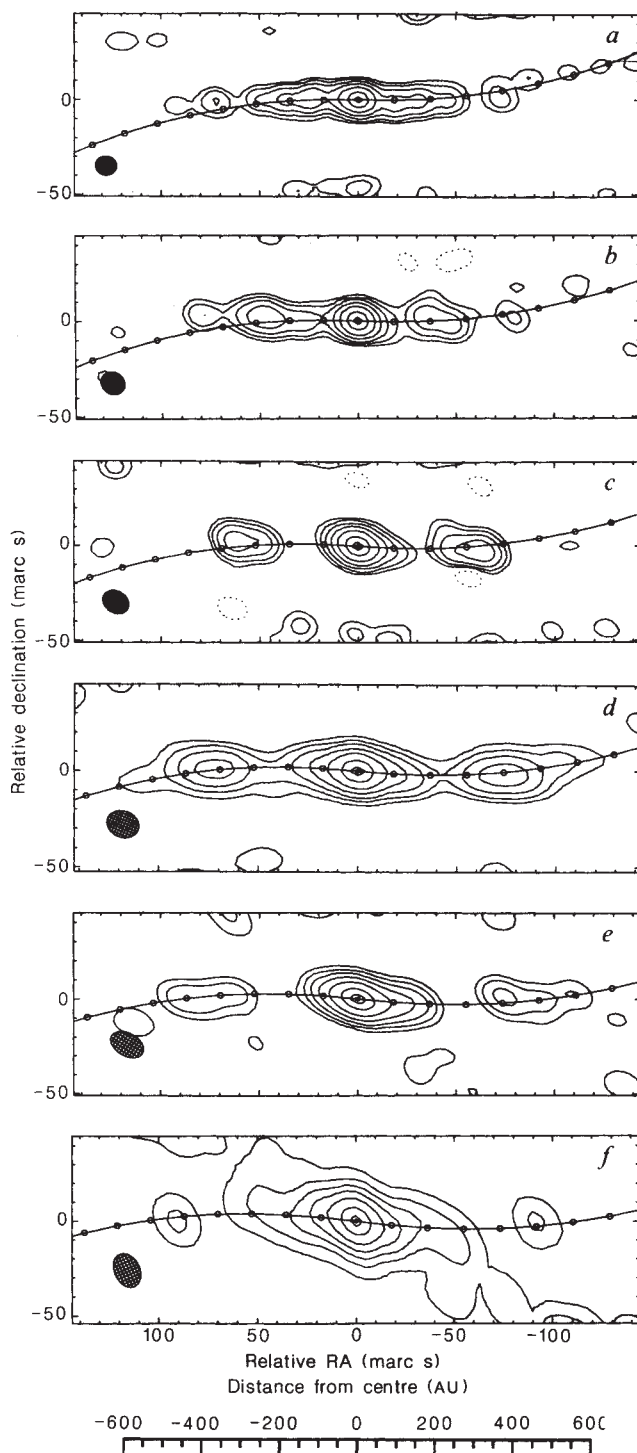


Fig. 1 VLBI maps of SS433 made at intervals of 2 days, starting on 17 May 1985. Contour levels have been drawn at $-2, 2, 4, 8, 16, 32, 64$ and 90% of the peak flux level in each map. The ellipse at lower left indicates the FWHM of the restoring beam. The curve in each map represents the locus of ejecta as predicted by the kinematic model parameters of Margon²; open circles mark intervals of two days. The linear scale in AU is also indicated.

of the speed predicted. Possible explanations include a low ejection velocity or a very deviant ejection angle. Such large deviations have never been seen in earlier radio or optical data², and in particular, are not seen in simultaneous optical spectra (R.C.V. *et al.*, in preparation). The symmetric flux density distribution in maps 3, 4 and 5 is hard to reconcile with an anomalous

Table 1 The Julian date (−2440000) of each of the six observations

		Effelsberg FRG	Jodrell Bank UK	Onsala Sweden	Medicina Italy	Westerbork Neth.
	Tsys (K)	65	125	80	70	70
	Sensit.(K Jy ^{−1})	1.4	0.076	0.06	0.17	0.35
	JD					
Map	−2440000					
1	6203.64	✓	✓	✓	—	
2	6205.63	✓	✓	✓	✓	✓
3	6207.62	✓	✓	✓	✓	✓
4	6209.63	✓	✓	✓	—	✓
5	6211.63	✓	✓	✓	✓	✓
6	6213.60	✓	✓	—	—	✓

A tick indicates the successful participation on that day of the station listed at the top of the column. The typical system temperature and sensitivity of each station is also listed.

Table 2 The flux density evolution of the core and of all the blobs we distinguish

	Flux density in each component (mJy)					Half life (days)	Proper motion (marc s per day)	Date of birth (JD)
	Map 1	Map 2	Map 3	Map 4	Map 5			
	3	2	—	—	—	47	3.8	6184.5
(West)	6	6	—	3	—	4	9.4	6199.0
	12	9	16	10	2	2	8.9	6200.3
	10	9	13	10	3	2	9.0	6201.6
	—	—	5	6	7	—	8.3	6202.5
(Core)	68	85	76	53	67	—	—	—
	—	—	—	4	6	—	9.7	6205.1
	—	15	10	9	5	4	8.4	6202.1
	15	29	15	8	2	2	8.8	6200.6
(East)	11	11	—	—	—	2	10.9	6200.2
	6	3	—	—	—	2	10.9	6198.9
	5	5	—	—	—	—	4.5	6187.4
	2	—	—	—	—	—	—	—

For fading blobs, the half-life in days has been calculated assuming a power law decay. The proper motion of the blobs is given in marc s per day, and the extrapolated date of birth in JD-2,440,000.

ejection angle of 30°, as we would then expect differential Doppler boosting²⁵ of a factor of 2. Explaining the deviation by enlarging the distance estimate of SS433 not only contradicts all other measurements, but also degrades the fit to the curved trace in our maps. We believe it is more plausible that there is a succession of individual blobs passing through a zone where they brighten, after which they rapidly decay.

The crosscuts shown in Fig. 2 display substructure. These 'plateaux' in Fig. 2 correspond to structures ('blobs') in Fig. 1 that are shaped like the synthesized beam shown. If we decompose each crosscut, we can find blobs which do, in fact, move in a way consistent with the kinematic model². Table 2 lists the flux densities of the core and of all the blobs evident in our maps. It is clear that a large proportion of the total flux density at 5 GHz is contained in the tabulated features. We also tabulate the velocities in the plane of the sky, and the extrapolated Julian date of ejection.

The behaviour of the blobs as apparent in Table 2 suggests the existence of a 'brightening zone', situated roughly 50 marc s from the centre ($\sim 4 \times 10^{15}$ cm at 5 kpc). That is to say, those knots which are located close to the centre at the start of our observing period first brighten, and then diminish in intensity. To the east, the brightening is most apparent between maps 1 and 2; to the west, between maps 2 and 3. Moreover, those blobs that only appear late in our observing session were still brightening at the time the observations were ended, while blobs already past the ~ 50 marc s mark in map 1 are only seen to decay. The brightest spots of emission along the precession locus are not stationary. This is to be expected if we observe a succession of blobs, rather than a continuous jet; after brightening, each blob will dominate the appearance of the maps for some

time while it moves ahead. This effect is enhanced apparently because some blobs happened to be rather brighter than their predecessors or successors.

Almost all of the radio emission in our maps lies on the locus predicted by the kinematic model². The apparently crooked central component in map 5 is, in our view, caused only by the convolution of the restoring beam with a structure that is somewhat more extended to the west than to the east. Only the outermost blob on either side in map 2 displays misalignment and deviant proper motion, which points to some unusual event. However, their flux density evolution is in accordance with the brightening zone hypothesis. We are unable to reject or confirm a connection of the deviant blobs with an earlier radio or optical flare.

Discussion

All previous investigators have concluded that the radio emission is due to synchrotron radiation, a conclusion we support. The radio structure is not resolved perpendicular to the beam, nor, as we argued in the previous section, along the beam. Assuming each blob to be a sphere, 15 marc s in diameter and at a distance of 5 kpc, taking minimum energy, a spectral index of -0.65 , cutoff frequencies of 10 MHz and 10 GHz, and a ratio of 100 between the total particle energy and the electron energy, we are able to derive an estimate for each blob of its luminosity (typically 5×10^{30} erg s^{−1}), brightness temperature (typically 2×10^6 K), energy (typically 7×10^{41} erg) and magnetic field (typically 8×10^{-2} gauss). The half-life of electrons emitting synchrotron radiation under such conditions is typically 8×10^5 years.

For those blobs in which a decline in flux density is seen, the

actual fading timescale is given assuming a power-law decay. Obviously, synchrotron losses cannot explain the rapid decline in flux density observed. It is more plausible that the knots undergo adiabatic expansion once they pass the brightening zone, located a few $\times 10^{15}$ cm from the core. The flux density decay seen would typically imply a 10% increase in radius.

Why would adiabatic expansion take place only at this distance? If we suppose that the stellar wind moves away from the companion star with the plausible speed of $2,000 \text{ km s}^{-1}$, it traverses $\sim 3 \times 10^{15}$ cm in one precession period. Thus if the wind gas has been disturbed (possibly even swept away) by the passing blobs, this gas could be fully replenished only in the inner precession cone during one precession period; before the next blob passes, replenishment at greater distances is expected to be less as it can take place only by diffusion of the more tenuous ambient matter at the sound speed. Therefore, it seems reasonable to attribute at least the adiabatic expansion beyond the brightening zone to the relative absence of confining material.

We are uncertain about the nature of the brightening zone. Many of the earlier maps²¹ show emission patches at a distance of 50 to 100 mas on either side of the core; and, again, a recent 18-cm map (I.F. *et al.*, in the press) contains knotty structure which supports the existence of a brightening zone. Hence, we must invoke a mechanism which always produces brightening at distances equal to a few hundred times the binary separation. There are a number of possible models, for instance, compression by an external medium, or optical depth effects. But in keeping with the evident ejection of gas in the form of discrete clouds, it is possible to explain the brightening without recourse to any external media. Brightening may occur after a new blob overtakes the bow shock of its predecessor. This hot bow shock, which forms at the leading edge of a cool clump of gas, points backward at the Mach angle. Because the ejection direction precesses, such overtaking will occur, even if the blobs move at the same speed. Several mechanisms may then cause radio-brightening: compression of the magnetic field, density enhancement, or *in situ* particle acceleration through turbulence. The time lapse between successive blobs is only a few per cent of the precession period. Thus, the time at which the bow shock is overtaken to first order depends only on the Mach angle. At a speed of $0.26 c$, the bow shock is traversed at about 4×10^{15} cm (50 mas) from the central source if the Mach number is 4.5.

An analysis of our simultaneous optical spectroscopy (R.C.V. *et al.*, in preparation) showed that at JD 2446202 there was a temporary deviation from the kinematic model² corresponding to a change in ejection angle of a few degrees, which is too small to be observed in our VLBI maps. Such a deviation may cause the blob to encounter more external material, leading to a more vigorous interaction with the ambient gas. Although the possible acceleration mechanism is obscure, such interaction could well lead to the brightening in the radio continuum which occurs at JD 2446201, and which may be related to the generation of somewhat brighter blobs closer to the core.

From the fact that the marked asymmetry observed in map 2 has disappeared in the next map, we conclude that such brightness differences are essentially transient and are due to the interaction between the blob and the ambient medium (maybe in the brightening zone), rather than being due to an intrinsic difference between the simultaneously ejected blobs.

We thank the staff at the EVN observatories who participated in these observations, W. Alef, H. Blaschke and the technical staff at the correlator centre in Bonn for their assistance, S. Bonsignori-Facondi, R. Fielder, K. J. Johnston and W. B. McAdam for performing radio flux density measurements and making the results available to us before publication, M. Watson for hospitality to one of us (R.C.V.) and for discussing Exosat results before publication. Stimulating discussions with E.P.J. van den Heuvel and A. R. Taylor are gratefully acknowledged. R.C.V. and I.F. acknowledge support from ZWO, the Netherlands Organization for the Advancement of Pure Research.

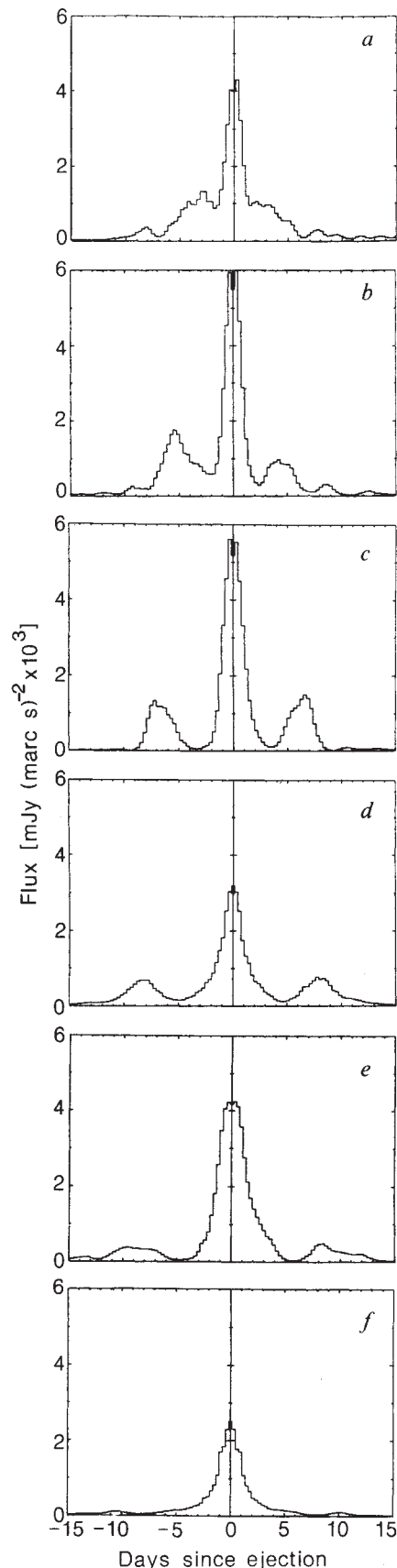


Fig. 2 Crosscuts through the maps of Fig. 1, along the kinematic model curve shown. The horizontal axis is labelled in time units (days) since the assumed ejection. All crosscuts have the same flux density scale, enabling one to follow the features from map to map.

Received 25 March; accepted 11 June 1987.

1. Abell, G. O. & Margon, B. *Nature* **279**, 701-703 (1979).
2. Margon, B. A. *Rev. Astr. Astrophys.* **22**, 507-536 (1984).
3. Anderson, S. F., Margon, B. & Grandi, S. A. *Astrophys. J.* **273**, 697-701 (1983).
4. Calvani, M. & Nobili, L. *Astrophys. Space Sci.* **79**, 387-395 (1981).
5. Gilmore, W. S. & Seaquist, E. R. *Astr. J.* **85**, 1486-1495 (1980).
6. Hjellming, R. M. & Johnston, K. J. *Nature* **290**, 100-107 (1981).
7. Gilmore, W. S. *et al. Astr. J.* **86**, 864-870 (1981).
8. Hjellming, R. M. & Johnston, K. J. *Astrophys. J.* **246**, L141-L145 (1981).
9. Hjellming, R. M. & Johnston, K. J. in *Radio Stars* (eds Hjellming, R. M. & Gibson, D. M.) 309-323 (Reidel, Dordrecht, 1985).
10. Spencer, R. E. *Nature* **282**, 483-484 (1979); *Mon. Not. R. astr. Soc.* **209**, 869-879 (1984).
11. Schilizzi, R. T. *et al. Astr. Astrophys.* **79**, L26-L27 (1979).
12. Walker, R. C. *et al. Astrophys. J.* **243**, 589-596 (1981).
13. Geldzahler, B. J., Downes, A. J. B. & Shaffer, D. B. *Astr. Astrophys.* **98**, 205-206 (1981).
14. Niell, A. E., Lockhart, T. G. & Preston, R. A. *Astrophys. J.* **250**, 248-253 (1981).
15. Niell, A. E. *et al. in IAU Symp.* **97**, 207-208 (1982).
16. Schilizzi, R. T. *et al. Nature* **290**, 318-320 (1981).
17. Schilizzi, R. T. *et al. IAU Symp.* **97**, 205-206 (1982).
18. Schilizzi, R. T. *et al. in Astrophysical Jets* (eds Ferrari, A. & Pacholczyk, A. G.) 157-159 (Reidel, Dordrecht, 1983).
19. Schilizzi, R. T., Romney, J. D. & Spencer, R. E. *IAU Symp.* **110**, 289-296 (1984).
20. Fejes, I. *Astr. Astrophys.* **168**, 69-71 (1986).
21. Romney, J. D. *et al. Astrophys. J.* (in the press).
22. Rogers, A. E. E. *et al. Science* **219**, 51-54 (1983).
23. Cohen, M. H. *et al. Astrophys. J.* **201**, 249-255 (1975).
24. Skilling, J. & Gull, S. F. in *Maximum Entropy and Bayesian Methods in Inverse Problems* (eds Smith, C. R. & Grandy, W. T. Jr) 83-132 (Reidel, Dordrecht, 1984).
25. Fejes, I. *Astr. Astrophys.* **166**, L23-L25 (1986).

Primary structure of the receptor for calcium channel blockers from skeletal muscle

Tsutomu Tanabe*, Hiroshi Takeshima*, Atsushi Mikami*, Veit Flockerzi*,
Hideo Takahashi*, Kenji Kangawa†, Masayasu Kojima†, Hisayuki Matsuo†,
Tadaaki Hirose‡ & Shosaku Numa*

* Departments of Medical Chemistry and Molecular Genetics, Kyoto University Faculty of Medicine, Kyoto 606, Japan

† Department of Biochemistry, Miyazaki Medical College, Miyazaki 889-16, Japan

‡ Pharmaceutical Institute, Keio University School of Medicine, Tokyo 160, Japan

The complete amino-acid sequence of the receptor for dihydropyridine calcium channel blockers from rabbit skeletal muscle is predicted by cloning and sequence analysis of DNA complementary to its messenger RNA. Structural and sequence similarities to the voltage-dependent sodium channel suggest that in the transverse tubule membrane of skeletal muscle the dihydropyridine receptor may act both as voltage sensor in excitation-contraction coupling and as a calcium channel.

VOLTAGE-dependent calcium channels in excitable membranes are essential for many cellular functions, including muscle contraction and secretory processes. Three types of calcium channel have been distinguished in neurons and two in muscle, by their electrophysiological and pharmacological properties¹. A variety of organic compounds, including 1,4-dihydropyridine derivatives, are known to modulate ion flux through slow L-type calcium channels². Calcium channels in skeletal muscle are localized in the transverse tubule (T-tubule) system of the cell membrane. The receptor for dihydropyridine calcium channel blockers, purified from rabbit skeletal muscle, has been reported to be composed of a large polypeptide of relative molecular mass (M_r) 130,000-170,000 (130K-170K) and one (M_r 30K-33K) or two smaller polypeptides (M_r 50K-56K and M_r 30K-33K)³⁻⁶. Recent immunoprecipitation studies with solubilized T-tubule membranes or with partially purified preparations indicate, however, that the rabbit skeletal muscle dihydropyridine (DHP) receptor comprises a single polypeptide of M_r 170K (ref. 6).

We have now deduced the primary structure of the large polypeptide of M_r 170K of the DHP receptor from rabbit skeletal muscle by cloning and sequencing DNA complementary to its mRNA. This polypeptide is homologous with the voltage-dependent sodium channel^{7,8} both in amino-acid sequence and in proposed transmembrane topology, containing four units of homology with characteristic putative transmembrane segments. The structure described has implications for the function of the DHP receptor.

Cloning of cDNA

The purified rabbit skeletal muscle DHP receptor preparation was subjected to phosphorylation by the catalytic subunit of cyclic-AMP-dependent protein kinase. SDS-polyacrylamide gel electrophoresis (PAGE) under reducing conditions showed that

the preparation comprised four major polypeptides of M_r 170K, 140K, 55K (triplet) and 33K (doublet) (Fig. 1A, lane a) and that the 170K polypeptide and the 55K polypeptide (doublet) were phosphorylated (Fig. 1A, lane b). The discrepancy between the M_r values of the phosphorylated large polypeptide reported from different laboratories^{5,9,10} may be attributable to its susceptibility to proteolysis⁶. Photoaffinity labelling with the dihydropyridine ³H-azidopine showed that the ligand was specifically incorporated into the 170K polypeptide (Fig. 1B) in agreement with a previous study with intact T-tubule membranes¹¹. We thus conclude that the 170K polypeptide is the DHP receptor. This conclusion is supported by recent immunoprecipitation experiments⁶.

The 170K polypeptide was separated from the other polypeptides by subjecting the purified DHP receptor preparation to gel permeation HPLC under denaturing conditions (Fig. 1C). The purified 170K polypeptide was digested with trypsin and the resulting peptides were fractionated by reverse-phase HPLC (Fig. 1D). Eleven fractions (I-XI) were collected and subjected to amino-acid sequence analysis (see Fig. 1 legend). Synthetic oligodeoxyribonucleotides were prepared on the basis of partial amino-acid sequences of the tryptic peptide X and used as a specific primer for reverse transcription and as a probe for selecting the resulting cDNA clones. The initial complementary DNA clone thus isolated was used as a probe for cloning longer cDNA sequences (see Fig. 4 legend).

Protein sequence

Figure 4 shows the 6,083-nucleotide sequence (excluding the poly(dA) tract) of the cDNA encoding the DHP receptor from rabbit skeletal muscle. There is an open reading frame that encodes a sequence of 1,873 amino acids (Fig. 4), including peptide sequences identical with all of those determined (amino-acid residues 169-174, 355-364, 365-382, 585-593, 689-698,

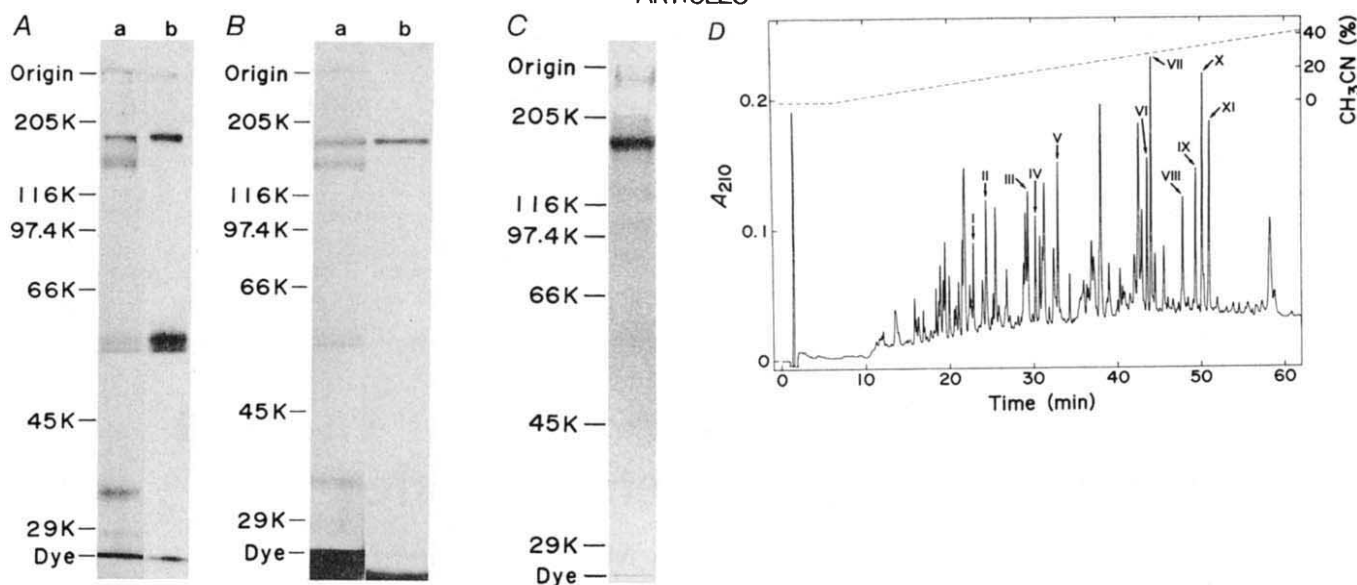


Fig. 1 Analysis of the DHP receptor protein purified from rabbit skeletal muscle. **A**, Silver stain (lane a) and autoradiogram (lane b) of SDS-PAGE of the phosphorylated DHP receptor; size markers (Sigma) are shown at left. **B**, Silver stain (lane a) and fluorogram (lane b) of SDS-PAGE of the DHP receptor subjected to photoaffinity labelling with ^3H -azidopine. **C**, Coomassie brilliant blue stain of SDS-PAGE of the purified 170K polypeptide. **D**, Reverse-phase HPLC of tryptic peptides derived from the 170K polypeptide. Solid line and left-hand axis, A_{210} ; broken line and right-hand axis, $\text{CH}_3\text{CN}(\%)$. The amino-acid sequences determined for eleven fractions (I–XI) corresponding to absorbance (A_{210}) peaks were as follows (in one-letter code): I, GLPDKTEEEK; II, GEGIPTTAK and TNPLAR; III, VLRPLR and EKQLEEDLR; IV, DGDPTQMELEPR; V, YDFEDTEVR and GYYYYYK; VI, LLDQVIPPIDDEVT; VII, DWSILGPHHL; VIII, FYATFLIQEHFR; IX, TGGLFGQVDTFLER; X, GYMSWITQGEVMDVEDLR; XI, AVPIPEA; fractions II, III and V were a mixture of two peptides.

Methods. The DHP receptor (0.66 μg), purified³ from rabbit skeletal muscle microsomes, was phosphorylated⁵ with the catalytic subunit of AMP-dependent protein kinase in the presence of 2 μM [^{32}P]ATP (78 Ci mmol^{-1}) (incubation at 37 °C for 10 min). The sample was boiled in gel loading buffer (20 mM Tris-HCl, pH 6.8, 4 M urea, 0.72 M 2-mercaptoethanol, 3% SDS and 1 mM phenylmethylsulphonyl fluoride (PMSF)) and electrophoresed on an SDS/7.5% polyacrylamide gel³⁴. The gel was silver-stained, dried and exposed to Fuji X-ray film at -70°C for 40 h. For photoaffinity labelling, 1.5 mg of the protein solubilized³ from T-tubules was incubated⁵ at 4 °C for 90 min at a protein concentration of 0.1 mg ml^{-1} in the presence of 20 nM racemic ^3H -azidopine³⁵ (45 Ci mmol^{-1}). The mixture was applied to a 2-ml column of wheat germ agglutinin (WGA)-agarose (Hohnen), and the bound protein was eluted with 0.3 M *N*-acetyl-D-glucosamine³; ~5% of the added ^3H -azidopine was recovered in the eluate. The eluate (~1.5 ml) was irradiated³⁵ for 2 min with a Sugiyamagen 30-W black-light lamp at 10 cm distance and purified further by sucrose density gradient centrifugation³. The purified material was dialysed against 5 mM Tris-HCl (pH 7.4) containing 0.05% SDS and 0.2 mM PMSF, lyophilized and dissolved in gel loading buffer. After boiling, the sample was divided into three equal portions (36,000 d.p.m. each) and electrophoresed as above. One lane was silver-stained, another lane impregnated with ENHANCE (NEN), dried and exposed to Fuji X-ray film at -70°C for 14 days, and the third lane sliced and counted for radioactivity. About 7% of the loaded radioactivity was incorporated into the 170K polypeptide. In a control experiment, in which incubation with ^3H -azidopine was carried out in the presence of 50 μM nifedipine, no significant amount of radioactivity was incorporated into the 170K polypeptide. To purify the 170K polypeptide, 2.4 mg of the DHP receptor purified³ from microsomes was boiled in gel loading buffer and fractionated by HPLC on a TSK 4000 column (7 $\times$ 600 mm)⁷. The 170K polypeptide obtained (~240 μg) was dialysed³⁶, and a portion (~2 μg) was boiled in gel loading buffer and electrophoresed as above. The peptides resulting from tryptic digestion of 162 μg of the dialysed material were fractionated by reverse-phase HPLC and sequenced⁷.

711–719, 770–776, 970–976, 977–988, 1,390–1,399, 1,505–1,520, 1,523–1,534, 1,601–1,614, 1,653–1,658). The nucleotide sequence surrounding the initiation codon agrees reasonably well with the consensus sequence¹². A nonsense codon (TAA, nucleotide positions –153 to –151) is found in-frame upstream of the initiation codon. The size of the DHP receptor mRNA has been estimated to be ~6,500 nucleotides by blot hybridization analysis⁸ of rabbit skeletal muscle poly(A)⁺ RNA with cDNA probes. From the cDNA sequence, we conclude that the rabbit skeletal muscle DHP receptor is composed of 1,873 amino-acid residues (including the initiating methionine), having a calculated M_r of 212,018. The difference between this value and the M_r estimated by SDS-PAGE (170K) may be attributable to proteolytic modification *in vivo* or during the purification of the protein, or alternatively to the inaccuracy inherent in the measurement of M_r of membrane proteins by the electrophoretic method¹³. The carboxy-terminal region, which has little secondary structure (see Fig. 2B), may be susceptible to proteolytic cleavage.

Structural characteristics

Similarity matrix analysis¹⁴ of the amino-acid sequence of the DHP receptor reveals the presence of four internal repeats (I–IV) that exhibit sequence homology (Fig. 2A, left): I, residues 38–337; II, residues 418–664; III, residues 786–1,068; and IV, residues 1,105–1,384. Among the repeats, 27(53)%, 30(51)%, 24(55)%, 27(51)%, 26(55)% and 21(48)% of the aligned positions are occupied by identical residues for the I–II, I–III, I–IV,

II–III, II–IV and III–IV pairs, respectively (where numbers in parentheses indicate values obtained when both identical and conservative residues are taken into account); a continuous stretch of gaps has been counted as one substitution regardless of its length. Furthermore, the DHP receptor is homologous in amino-acid sequence with the sodium channel, particularly in the regions comprising the four internal repeats, as shown by similarity matrix analysis (Fig. 2A, right) and sequence alignment (Fig. 4). Between the corresponding repeats of the DHP receptor and rat sodium channel II, 32(62)%, 35(59)%, 37(60)% and 32(61)% of the aligned positions are occupied by identical (and identical or conservative) residues for repeats I, II, III and IV, respectively. The degree of overall sequence similarity between the two proteins is 29(55)%. This observation suggests that the genes encoding the DHP receptor and sodium channels arose from a common ancestor. Thus it is inferred that voltage-dependent ionic channels represent a family of evolutionarily and structurally related gene products.

The DHP receptor has a hydropathicity profile similar to that of the sodium channel^{7,8} (Fig. 2B). Each internal repeat has five hydrophobic segments (S1, S2, S3, S5 and S6) and one positively charged segment (S4), all of which exhibit predicted secondary structure (Fig. 2B). The six segments share characteristic structural features with the corresponding segments of the sodium channel^{7,8} (Fig. 3). The DHP receptor, like the sodium channel, does not possess a hydrophobic amino-terminal sequence indicative of the signal sequence. Thus it seems reasonable to

Fig. 2 **A**, Similarity matrix analysis and **B**, hydrophobicity profile of the DHP receptor. In **A**, the amino-acid sequence of the DHP receptor (ordinate) is compared with the sequences of itself (abscissa, left) and rat sodium channel II (abscissa, right) using a computer program to generate diagonal lines indicating segments of 25 residues that show sequence similarity (scores ≥ 70)¹⁴; the computer program includes conservative substitutions as well as identities. Bars, internal repeats I-IV of the DHP receptor and sodium channel II. In **B**, the averaged hydrophobicity index⁴¹ of a nonadecapeptide composed of amino-acid residues $i-9$ to $i+9$ is plotted against i , the amino-acid number. The positions of the predicted structures of α -helix and/or β -sheet⁴² that have a length of 10 or more residues are shown by open boxes below. The positions of positively charged residues (Lys and Arg) and negatively charged residues (Asp and Glu) are indicated by upward (+) and downward (-) vertical lines, respectively. Bars above, positions of repeats I-IV; the positions of segments S1-S6 in each repeat are also indicated by differently filled boxes.

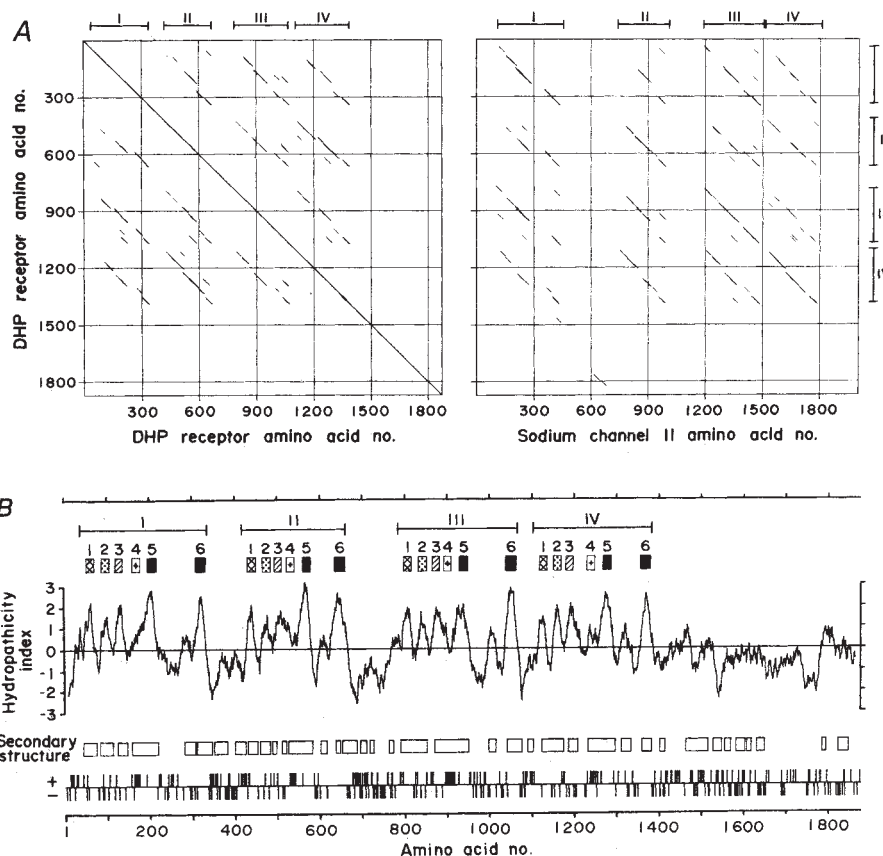
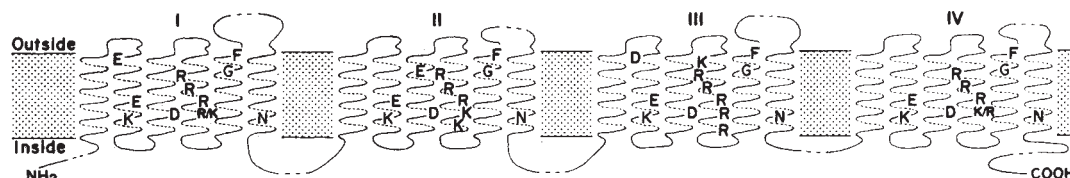


Fig. 3 Schematic representation of structural characteristics common to the DHP receptor and the voltage-dependent sodium channel. The four units of homology spanning the membrane, which are assumed to surround the ionic channel, are displayed linearly. The putative transmembrane α -helical segments S1-S6 (from left to right) in each repeat (I-IV) are illustrated. The Arg and/or Lys residues in segment S4 conserved in the individual repeats of the DHP receptor and the three sodium channels of known sequence^{7,8} are shown in one-letter code. The residues in the other segments conserved in all repeats of the DHP receptor and the three sodium channels are also shown and are as follows (amino-acid numbers indicating those of the DHP receptor): Glu residues 100, 478, 846 and 1,164 and Lys residues 104, 482, 850 and 1,168 in segment S2; Asp residues 126, 500, 872 and 1,186 in segment S3; Gly residues 214, 577, 946 and 1,285 and Phe residues 218, 581, 950 and 1,289 in segment S5; Asn residues 327, 654, 1,058 and 1,374 in segment S6. In addition, Glu 90 and Asp 836 in segment S2 conserved in repeats I and III, respectively, and Glu 510 in segment S3 conserved in repeat II of the four proteins are indicated.



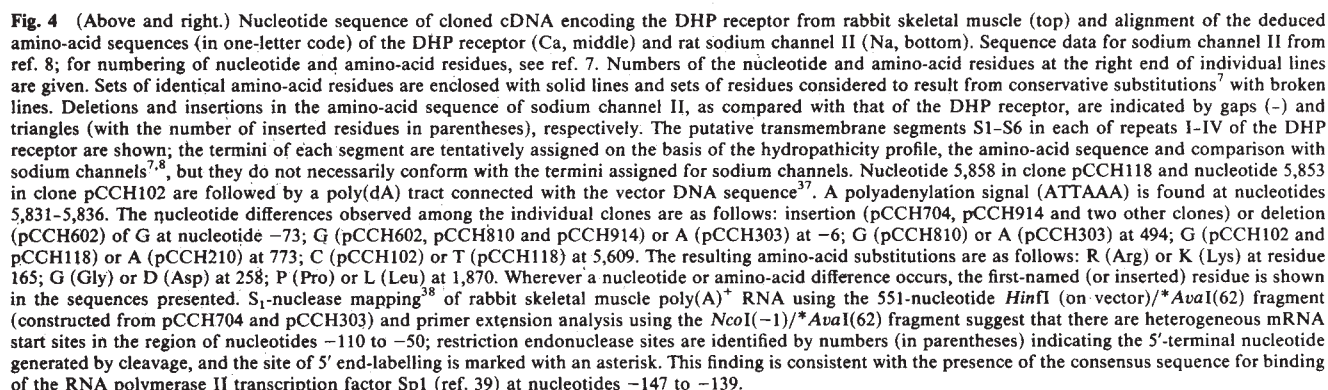
assume that the DHP receptor molecule has a transmembrane topology similar to that proposed for the sodium channel^{7,8} in that the four repeated units of homology, each containing the six presumably α -helical membrane-spanning segments S1-S6, are oriented in a pseudosymmetric fashion across the membrane and in that the amino- and carboxy-termini are on the cytoplasmic side of the membrane (Fig. 3). This model is consistent with two of the five potential *N*-glycosylation sites¹⁵ (Asn 79 and Asn 257) and all the potential cAMP-dependent phosphorylation sites¹⁶ (Ser residues 687, 1,502, 1,575, 1,757, 1,772 and 1,854 and Thr 1,552) being located on the extracellular and on the cytoplasmic side, respectively.

Segment S4 in every repeat of the DHP receptor contains five or six Arg or Lys residues at every third position (except for the basic residue closest to the carboxyl end of segment S4 in repeats III and IV residing at the fourth position), with mostly nonpolar residues intervening between the basic residues. Such a structure is highly conserved in segment S4 of all sodium channels of known sequence^{7,8} (Fig. 3). It seems probable that the positive charges in this segment, many of which presumably form dipoles, represent the voltage sensor^{7,8}. The conserved

charged residues in segments S2 and S3 of sodium channels^{7,8} are also shared by the DHP receptor (Fig. 3). Note also that all repeats of the DHP receptor as well as of sodium channels contain a Gly and a Phe at equivalent positions in segment S5 and an Asn at an equivalent position in segment S6 (Fig. 3).

Functional implications

Previous cDNA expression studies suggest that the function of the sodium channel can be manifested by the large polypeptide of $M_r \sim 260K$ alone^{17,18}. Thus the striking structural similarity found between the DHP receptor and the sodium channel may imply that the DHP receptor itself is a calcium channel. This notion is consistent with the finding that reconstitution of the purified skeletal muscle T-tubular DHP receptor into phospholipid bilayer membranes or vesicles yields a functional calcium channel^{19,20}. In view of the diameter of the calcium channel being $\sim 6 \text{ \AA}$ (ref. 21), one or two helical segments from each repeat would take part in constructing the inner wall of the channel at the narrowest point. Under physiological conditions, the calcium channel is highly selective for Ca^{2+} , and this selectivity has been proposed as being derived from high-affinity



the membrane, bears some resemblance to the EF-hand, and the sequence of residues 155–164, assigned mostly to the extracellular side, is similar to a portion of the consensus sequence repeat of Ca^{2+} -dependent membrane-binding proteins. It is attractive to hypothesize that the two Glu residues (residues 87 and 90 in repeat I) and the three Asp residues (residues 465,

AAGGCGAAGTCTTCAGTCGCAACGACCTATCCAAAGATGACAGAAGGAGTGCAGGGGGTACTACTTGTGTGACGAAGCGGGGACCCACGACGATGAGCTGCGCCCGCCGACGTGATACACAACTGCCTCCACTTTCGACAACTG
Ca K G K F F S C N D L S K M T T E E D C R G V Y Y V Y K D G D P L T Q M E L R P R Q W I H N D F H F D N V 1000
Na A G K F F H C I N Y - - - T T I G E D V S V V N Y Y S C Q A P L T E S N T A R L W K N V K V N F D N V 1408

CTGTGGCCGATGATGCTGCTCTACCGGTGTCACCTTCGAGGATG6CCGACGCTGCTGACAGGGCCATAGACTCCCAACGAGGAGGACATGGCCCGGTTCACACAAACGAGTGAGATG6CCATCTTCTCATCATCATCATCTC
Ca L S A M M S L L F T V S Y F E G W P Q L L V R A I T D S N E E D M G P V Y N N R V E M A I Y F F V I I I Y I I 3150
Na L G L Y I S L L Q V A T F K G W M Q D I M Y A I A V D S R N V L L Q P K Y E D N L Y M Y L Y F F I I I Y I I 1458

CTCATTCGCTTCTCATGATGAACATCTTTGTGGCTTGTGCATCGTCACCTTCGAGGAGCAGGGGAGACAGAGTACAAGAACTGCGAGTGGACAAGAAGCAGCGCCAGTGTGTGAGATATG6CCTGAAGGCCCGCCCATCTCGGTGC
Ca L A F F M M N I F V G F V I V T F Q E Q G E T E Y K N C E L L D K N Q R Q V Q Y A L K A R P L R C 3300
Na F S S F F T L N L F I G V I I D N F N Q Q K K G G Q D I F M T E E Q K R Q Y N A M K K L G S K K I 1103

TACATCTCCGCAAGCACTACGACCAAGGTGTGGTACGTCGTCACTCTCTCTACTTTGAATACCTGATGTGCGCCTCATCATGCTCAACACCATCTGCTGGGCATGAGCAGTACCACCAGTCGGAGGAGATGAACACATCTGC
Ca Y I P X N P P Y Q I Y Q V W F V V T S S Y F E Y L M F A L I L N I T C L G M Q H Y H Q S E M M N H I S 3450
Na P R P A N K K F Q G M V F D F V T T K Q Y F D I S I M L I L I C L N M V T M M V E T D D Q S Q E M T N I L 1163

GACATCTCTCACTGGCTTCCACCATCATCTTCACACTGGAGATGATCTCAAGCTCTTGGCTTCAAGGCGAGGGGCTATTTCGGAGACCCCTGGAATGTGTCGACTTCTGATGCTCATCGGAGCATTCATGACGTCACTCCAGC
Ca Y I L N V A P T I I F I I E M L K L L A F K A R G Y F T G P W N V I F D F L I V I G S T I D V I L S 3600
Na Y I I N L V A P T I I V L F T G E C M V L K L L I S L R I H Y Y F I T G W N I I F D F V I V I L S Y V G M F I A 1202

GAGATCGACACTTTCGTGGCTCCAGCGGGGGACTGATTGCTGGGTGGCGCTGCGGGGAAGCTTGACCCAGACGAGAGCGCCGCTCTCTCAAGTGCTTCTCCGCTGCTTCGGGCTGATGAGCTGATCAAGCTGCTGAGTGGGGC
Ca E I D T F L A S S G G L Y C L G G G C G N V D P D E S A R I S S A F F R I L F R V M R L L K L L S R A 3750
Na E I I E K Y F V S P L L F - - - - - R V I R L A R I G R I L R L L K G A K G I 1245

GAGGCGTGGCAGCGTCTGTGGAGCTTCAAGTCTTCAGGCGCTGCCCTACGTGGCTGTCTCATGCTCATGCTGTCTTCATCTACGCGCTCATCGGATGCGAGATGTTTGGAAAGATGCGCTGGTGACGGGACCGACATC
Ca E G V R - - T L L F A T F I K S F Q A L P V I A L L I V M L F F F I Y A Y I G M Q M F A Y K I A L V D G T Q I 3900
Na R - - - T L L F W T F M S L F A L P N I G L L L F L V M F I Y A I F G M S N F A Y Y V K R E V G - - I 1631

AATCCGCAACACAACCTTCAGACCTTCCGCGAGGCGTGTGCTGCTCTCAAGTGTGGCAGGAGGAGGCTGGGACAGATCTGCTGGCTCGACGTACGCGAAGTGTGCGACCCAGAGTACAGACTACGCCCGGGCGAGGATC
Ca N R N N N F Q T F P Q A V L F R C A T T G E A N W Q E I L L A C S Y G K L C D P E S D Y A P G E E Y 4050
Na D M D N F N F E T F G N S M I C L F L Q I T T S A G W D G L L L A P I L N S G P C D P E K D N P G S S K G 1344

ACGTGTGGCACAACCTTCGCTACTACTACTTCATCAGCTTCTACATGCTCTGCGCTTCTGATCATCAACTCTTCTGGCTGTCTCATGACAACCTTGTACCTGACACCGGACTGGTTCATCTGGGCGCTTCCACCTTGGAG
Ca T C G T N F A Y Y V F Y S Y F Y M L C A F L I I N L F V A V I M D N F D Y L T R D W S I L G P H D L D 4200
Na D C G P S V G I F F F V F S Y I I S F L V V V N M Y I A V I L E N F S V A T E E S A P L S D H D F 1796

GAGTCTCAAGCTTCTGGCAGAGTATGACCCAGAGGCCAGAGGCGAATCAAGCAGCTGGAGCTGTGACCTGTGAGAAGGATCCAGCCCTCTGGGCTTCGGGAAGTCTGTCCACACGGGTGGCTTAAGCGCTGGTGGG
Ca E F K A I W A E V Y D P E A K G R I K H L D V V T L R R I T Q P P L G F G K F C P H R V A C K R L V I G 4350
Na M F Y E V W E K F D P D A T K Q F I E F C K I S D F A A A L D P P L L I A I A - - P N K R V - - - Q L L I A 1841

ATGAACATGCCCTGAACAGTACGGCAGGTCACCTTCAATGACGAGCTCTTGGCTGTGGTGGCAGGCGCTCAAGATCAAGACAGAAGGTAACCTTGAAGAGGCAACGAGGAGCTGAGGCGCATCATCAAGAAGTCTGGAAGAGA
Ca M N D L P L N S D G T V F N A I T L F A L V R T A L K T K T E G N F E Q A N E E L R A I I K K I W K R 4500
Na M D L P L M V S G D R I V C L D I L F A F T K I - - R V L G E S G E M D A - - - - - 1874

ACCAGATGAAGCTGTGGACAGGTATCTCTCCATAGGAGTACGAGGAGTGCCTGGGAGAGTCTACGCAACATCTCATCAGGAGCATTCGCGAGTTCATGAAGCGCCAGGAGAATATTATGGTATCGGCCCAAGAA
Ca T S M K L L D Q V I P P I G D D E V Y V G K F Y A T F L I Q E H F R K I M K R Q E E V Y Y G Y R P P K K 4650
Na T S M K L L D Q V I P P I G D D E V Y V G K F Y A T F L I Q E H F R K I M K R Q E E V Y Y G Y R P P K K 1550

GACACCGTGCAGATTCAGGCTGGCTGCGACCATAGAGAGGAGGCGCCCTTGAGATGCGCCGACCATCTCAGGAGACTGACCCCGAGAGGAGTTCGAGAGAGCCATGTTGAGGCTGCGATGAGGAGGATCTTCGAGAG
Ca D T G L Q I A G L R I G G E T A P E I R T T I S G D T A E E L E A M F E R A M F E R 1600
Na V S Y E E T T T L L R K G E E S A I I L Q R A Y R Y R L L A Q K Y K K V S S I I Y K K D K G K E D E G 1942

ACGGAGGCGCTGTTGGCAGGTGACACCTTCTGGAAGGACCACTTCTGCCCGCGGTGATGGCCCAAGAGACGCTCGAGTGTGAGATGAGAACTGGAAGAGCTGAGTGGCTGTCTTCTGAGAGCTTCTTCAGAGT
Ca T T G L F G Q A G L R I T T I P P V Y A N Q R P Q A E I K A E I L E S P A M F E R D K P Q D 1650
Na T F I K E D I T T T P L E K T D V T P S T T S Y D S V T K P L E K F K F E K D K - S E K E D K G K D 1999

GCAAGAACACACCTCTCGCTGTGCCAATCAACCAAGCCAAATGTTGCTATGGCAGCAGCAACCATGACCAACAGATGTTTCCAGCGTCCATGTGAAGGAGTTCCTGGGAGGCGGAGACCTCCGCTGCGGA
Ca A R T N P L A R T A N T A N N V A Y G N S N H S N C A G S V S H C E R A G E A T P A A G 1700
Na I R - E S K - - - - - 2005

CGAGGAGCCTCGACCATCCACAGGCGCTGGGACCTCAGCAGCGCTGTGCTGGAAGAACTGAATGGCAGCTGGTCCAGCGGGGATGCCCATCAACAGGCGACCTCTGCGCCCTGCCAGCGCTAGCAGGATCCCGAGAG
Ca R G C L S H S R A L G P H S K P A C G A K L N G L V Q P G M P I N Q A P P A P C Q Q P S T D P P E 5250
Na R G C L S H S R A L G P H S K P A C G A K L N G L V Q P G M P I N Q A P P A P C Q Q P S T D P P E 1750

CGCGGGCAGAGGAGGACCTCCCTGACAGGCTCTCTGCAAGCAGAGACCCCGAGGAGGAGCTCGAGGGGAGCACCCTCCAGGCGCGGCTCTGCTGATCAGCTGTGATCGAAGGCTGTGTTGCGAGGGGCTGGACACCTTG
Ca R G Q R R T S L T G S L Q D E A P Q R R S S E G S T P R R P A P A T A L L I Q E A L V R G G L D T L 5400
Na R G Q R R T S L T G S L Q D E A P Q R R S S E G S T P R R P A P A T A L L I Q E A L V R G G L D T L 1800

GCAGCTGATGCTGCTCTGACGAGCAACAGCCAGGCGCTGGCAGAGCGCTGTGATGGAAGCGAGGAGAGTGAAGTGCAGCCACAGAGCTCAAGAGCGGAGAGTCTCGGAGGATGCGCAGTGTCCCGGAGGCTGCGACCTGAGC
Ca A A D A G F V T A T S Q A L A D A C O M E P E E V E V A A T E L L K A R E S V Q G M A S V P G S L S 5550
Na A A D A G F V T A T S Q A L A D A C O M E P E E V E V A A T E L L K A R E S V Q G M A S V P G S L S 1850

CGAGGCTTCCCTTGGCAGCTTGACACAGGCTCAGGCGCTCCAGCAAAACCTTATCTCCCGAGGCGGATGCTGTGTTGCTCAGCATGACCAAGCGAGGAGCAAGTGGCTGCAAGAGCTCAGGCTTGCATGGCAGCTCCCTCTG
Ca R R S L G S L D Q V G S Q E T L I P P R P 5700
Na R R S L G S L D Q V G S Q E T L I P P R P 1800

TCTCAGCCCTCTCTGAGTGGGGGGTCTGGAGCGCAGCAGGAGCGAGGAGCTCCCTGGGCGAGCAGGAGCATGTATTAAGCCATCCAGAAAGGCTGGTCACTGCCATCCCGCAGGACATTAAGATCTCTAGCTGTCT
Ca GGCATGG-----3' 5850
Na GGCATGG-----3' 5850

Methods. For details of cDNA cloning procedures, see ref. 7; screening was carried out by hybridization at 50 °C with 5'-³²P-labelled probes unless otherwise specified. The cDNA resulting from primer extension with 3.2 nmol of 5'-AC^ΔTCATAC^ΔTC-3' (synthesized as two pools containing A/T or G/C for N₃; corresponding to the sequence EVMDV of the tryptic peptide X) and 200 μg of rabbit skeletal muscle poly(A)⁺ RNA was cloned in pBR322. Screening of ~6 × 10⁵ transformants by hybridization at 33 °C with an equimolar mixture of three synthetic probes, 5'-ATCCAGACAT^ΔTA-3' as two pools and 5'-ATCCA^ΔCTCAT^ΔTA-3' (corresponding to the sequence YMSWI of the peptide X), yielded one clone, pCCH7 (1,007-1,119) harbouring a coding sequence for part of the peptide X; numbers in parentheses indicate the nucleotide residues carried by individual clones. A cDNA library³⁷, constructed with 4.2 μg of the vector-primer DNA and 7.8 μg of poly(A)⁺ RNA, was screened using the *EcoRI*(1,007)/*Sau3AI*(1,107) fragment to yield three positive clones, including pCCH102 (727-5,853) and pCCH118 (727-5,858; nucleotides 1,009-4,490 were not sequenced), from ~1.7 × 10⁶ transformants. Further screening of clones derived from the initial primer extension (~1 × 10⁶ transformants) using the *AluI*(748)/*NcoI*(870) fragment gave four positive clones including pCCH210 (527-876). Another synthetic primer (1 nmol) complementary to nucleotides 649-665 was elongated using 200 μg of poly(A)⁺ RNA. The resulting clones were screened three times (~2 × 10⁵ transformants each time): the *BanI*(535)/*SacI*(651) fragment yielded ~300 positive clones including pCCH303 (-69 to 663), the *RsaI*(272)/*BanI*(535) fragment ~120 positive clones including pCCH810 (-69 to 616), and the *SacII*(-41)/*SacII*(86) fragment 19 positive clones including pCCH914 (-147 to 216). Elongation of a third synthetic primer (1 nmol) complementary to nucleotides 2-18 (using 200 μg of poly(A)⁺ RNA) and screening of the resulting clones (~3 × 10⁵ transformants) with the *SacII*(-41)/*NcoI*(-1) fragment yielded 36 positive clones including pCCH602 (-164 to 15). The same filters were re-used for screening with the *AluI*(-159)/*AluI*(-99) fragment, giving two positive clones including pCCH704 (-225 to -38). Sequencing⁴⁰ of the cDNA was carried out on both strands.

836 and 1,151 in repeats II, III and IV, respectively) are involved in Ca^{2+} binding. These negatively charged residues are assigned to the vicinity of the extracellular end of the putative transmembrane segment S2 and are assumed to be located on an equivalent side of the four transmembrane α -helices.

On the other hand, Ca^{2+} entry through the surface membrane

is thought to be unimportant for contraction of skeletal muscle²⁷. Furthermore, it has been reported that less than a few per cent of the DHP-binding sites in skeletal muscle represent functional calcium channels²⁸. Thus the possibility arises that the T-tubular DHP receptor has a function other than that of a calcium channel. A plausible role for the DHP receptor is to act as the

voltage sensor involved in excitation-contraction (E-C) coupling to release Ca^{2+} from the sarcoplasmic reticulum (SR). This concept is supported by the finding that DHP inhibits charge movement in the T-tubule membrane and Ca^{2+} release from the SR in parallel and that this effect has a membrane voltage dependence analogous to that of DHP binding²⁹. Moreover, charge movement becomes significant at more negative voltages than Ca^{2+} release, whereas both seem to saturate at similar positive membrane potentials³⁰.

This observation favours a gating system in which charge movement is generated in more than one step but in which only the final step activates Ca^{2+} release³⁰. Such a gating mechanism is consistent with the presence of four putative voltage sensor segments (S4) in a single DHP receptor molecule, as discussed for the sodium channel^{7,8}. In view of the structural characteristics of the skeletal muscle DHP receptor, it is tempting to hypothesize that this molecule has a dual function in T-tubules. It may

primarily function as the voltage sensor for E-C coupling and may also act as a calcium channel to ensure long-term Ca^{2+} replenishment. This notion is supported by the fact that both a decrease in DHP-binding sites³¹ and a lack of the slow calcium current³² are observed in mice with muscular dysgenesis, a mutation that selectively disrupts E-C coupling³³. It seems possible that the carboxy-terminal region of the DHP receptor, which is larger than that of the sodium channel, is involved in E-C coupling.

We thank Drs Paul Berg and Hiroto Okayama for their high-efficiency cloning system and Dr Takashi Miyata and Mr Kei-ichi Kuma for computer analysis. Supported in part by grants from the Ministry of Education, Science and Culture of Japan, the Mitsubishi Foundation and the Japanese Foundation of Metabolism and Diseases and by fellowships from the Japan Society for the Promotion of Science (T.T. and V.F.) and the Feodor Lynen Program (V.F.).

Received 12 May; accepted 24 June 1987.

- Reuter, H. *Nature* **316**, 391 (1985).
- Reuter, H., Porzig, H., Kokubun, S. & Prod'homme, B. *Trends Neurosci.* **8**, 396-400 (1985).
- Curtis, B. M. & Catterall, W. A. *Biochemistry* **23**, 2113-2118 (1984).
- Borsotto, M., Barhanin, J., Norman, R. I. & Lazdunski, M. *Biochem. biophys. Res. Commun.* **122**, 1357-1366 (1984).
- Flockerzi, V., Oeken, H.-J. & Hofmann, F. *Eur. J. Biochem.* **161**, 217-224 (1986).
- Vandaele, S., Fosset, M., Galizzi, J.-P. & Lazdunski, M. *Biochemistry* **26**, 5-9 (1987).
- Noda, M. *et al.* *Nature* **312**, 121-127 (1984).
- Noda, M. *et al.* *Nature* **320**, 188-192 (1986).
- Curtis, B. M. & Catterall, W. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2528-2532 (1985).
- Hosey, M. M., Borsotto, M. & Lazdunski, M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3733-3737 (1986).
- Galizzi, J.-P., Borsotto, M., Barhanin, J., Fosset, M. & Lazdunski, M. *J. biol. Chem.* **261**, 1393-1397 (1986).
- Kozak, M. *Nucleic Acids Res.* **12**, 857-872 (1984).
- Grefrath, S. P. & Reynolds, J. A. *Proc. natn. Acad. Sci. U.S.A.* **71**, 3913-3916 (1974).
- Toh, H., Hayashida, H. & Miyata, T. *Nature* **305**, 827-829 (1983).
- Hubbard, S. C. & Ivatt, R. J. *A Rev. Biochem.* **50**, 555-583 (1981).
- Krebs, E. G. & Beavo, J. A. *A Rev. Biochem.* **48**, 923-959 (1979).
- Noda, M. *et al.* *Nature* **322**, 826-828 (1986).
- Stühmer, W., Methfessel, C., Sakmann, B., Noda, M. & Numa, S. *Eur. Biophys. J.* **14**, 131-138 (1987).
- Flockerzi, V. *et al.* *Nature* **323**, 66-68 (1986).
- Curtis, B. M. & Catterall, W. A. *Biochemistry* **25**, 3077-3083 (1986).
- McCleskey, E. W. & Almers, W. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7149-7153 (1985).
- Hess, P. & Tsien, R. W. *Nature* **309**, 453-456 (1984).
- Almers, W. & McCleskey, E. W. *J. Physiol., Lond.* **353**, 585-608 (1984).
- Kretzinger, R. H. A. *Rev. Biochem.* **45**, 239-266 (1976).
- Stuart, D. I. *et al.* *Nature* **324**, 84-87 (1986).
- Geisow, M. J., Fritzsche, U., Hexham, J. M., Dash, B. & Johnson, T. *Nature* **320**, 636-638 (1986).
- Miller, R. J. & Freedman, S. B. *Life Sci.* **34**, 1205-1221 (1984).
- Schwartz, L. M., McCleskey, E. W. & Almers, W. *Nature* **314**, 747-751 (1985).
- Rios, E. & Brum, G. *Nature* **325**, 717-720 (1987).
- Melzer, W., Schneider, M. F., Simon, B. J. & Szucs, G. *J. Physiol., Lond.* **373**, 481-511 (1986).
- Pincon-Raymond, M., Rieger, F., Fosset, M. & Lazdunski, M. *Dev. Biol.* **112**, 458-466 (1985).
- Beam, K. G., Knudson, C. M. & Powell, J. A. *Nature* **320**, 168-170 (1986).
- Powell, J. A. & Fambrough, D. M. *J. cell. Physiol.* **82**, 21-38 (1973).
- Laemmli, U. K. *Nature* **227**, 680-685 (1970).
- Ferry, D. R., Rombus, M., Goll, A. & Glossmann, H. *FEBS Lett.* **169**, 112-118 (1984).
- Kawakami, K. *et al.* *Nature* **316**, 733-736 (1985).
- Okayama, H. & Berg, P. *Molec. cell. Biol.* **2**, 161-170 (1982).
- Berk, A. J. & Sharp, P. A. *Cell* **12**, 721-732 (1977).
- Kadonaga, J. T., Jones, K. A. & Tjian, R. *Trends biochem. Sci.* **11**, 20-23 (1986).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
- Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105-132 (1982).
- Chou, P. Y. & Fasman, G. D. A. *Rev. Biochem.* **47**, 251-276 (1978).

LETTERS TO NATURE

The progenitor of SN1987A

R. Gilmozzi*, A. Cassatella*, J. Clavel*,
C. Fransson†, R. Gonzalez*, C. Gry*,
N. Panagia‡, A. Talavera* & W. Wamsteker*

* ESA IUE Observatory, aptdo 54065, Madrid, Spain

† Stockholm Observatory, S-133 00 Saltsjöbaden, Sweden

‡ Space Telescope Science Institute, Baltimore, USA, and
Istituto di Astronomia, Università di Catania, Italy

One of the reasons why the bright supernova SN1987A is of such importance is the fact that its proximity would allow for the first time, the identification of a supernova progenitor star. Attempts to establish which star was the progenitor have led to uncertainties, however. Here we use spatial and spectroscopic information from International Ultraviolet Explorer (IUE) spectra in the range 1,150-1,950 Å to demonstrate that Sk-69 202, the star coinciding positionally with the Large Magellanic Cloud (LMC) supernova, has disappeared from sight. Two weaker sources, named star 2 and star 3 in the astrometric analysis of the image of Sk-69 202 before the supernova outburst, are still present in the ultraviolet spectra. We isolate their spectra and give a spectral classification. We conclude that Sk-69 202 is the progenitor of SN1987A.

Very soon after its discovery, the positional coincidence between the star Sk-69 202 and the supernova gave rise to the suggestion that this star was the progenitor of SN1987A^{1,2}. A

reanalysis of previous plates of the region of the supernova allowed one to determine that three stars were actually present within 3 arcs from the supernova position³ (see Fig. 1c). The very close positional agreement between the supernova and star 1 (B3 Ia; $V = 12.36$; $(B - V) = 0.04$) strongly suggested that this star was the progenitor, while the position of star 2 (ref. 3; N. R. Walborn *et al.* preprint, 1987) (3 arcs at position angle (PA) 315°, $V = 15.30$; $(B - V) = 0.2 \pm 0.4$) made it clear that star 2 could be immediately excluded. The presence of a third star much closer to star 1 was, however, indicated. The detailed analysis (N. R. Walborn *et al.* preprint, 1987) gave 1.5 arcs at PA 120°, $V = 15.7$ and $(B - V) = 0.15 \pm 0.4$ for star 3. These results appeared to suggest that star 1 had actually been the progenitor. But various ultraviolet results from the IUE satellite did not seem to be consistent with such a conclusion. After the rapid decay in the ultraviolet, the far ultraviolet spectrum showed a residual stellar spectrum of an early B type, and this was taken to support the idea that star 1 was still present⁴. In later reports⁵ it was argued from the multiplicity seen in the IUE spatially resolved spectra that the spectrum of star 2 was also still visible. In an earlier paper⁶ the authors argued that the spatial information from IUE spectra was in better agreement with star 1 having actually disappeared and with the two visible spectra belonging to stars 2 and 3. This was later confirmed by other IUE observers⁷. As it will probably be some time before the optical observations will allow us to clarify this question and in view of the importance of the identification of the progenitor, here we present a new detailed analysis of some IUE spatially resolved spectra together with new ultraviolet observations showing unambiguously that star 1 has disappeared and

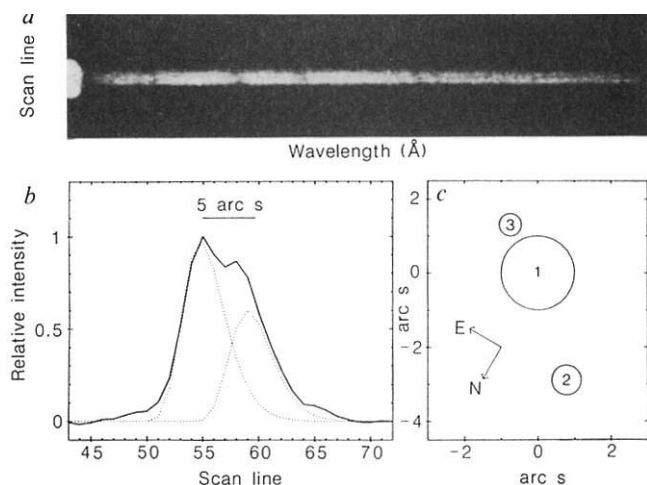


Fig. 1 *a*, Composite line-by-line spectrum between 1,210 and 1,510 Å, obtained after the supernova had faded. The two-point source spectra shown to correspond to stars 2 and 3 (see *c*) can be easily distinguished. *b*, Cross-cut through the dispersion (1,300–1,500 Å). The dotted lines represent the PSF fits. A 5 arc s bar is drawn for reference. See text for details. *c*, The configuration of Sk-69 202 as seen in the large aperture on 3 March (PA = 330°).

that the remaining ultraviolet flux in the short wavelength range is contributed entirely by stars 2 and 3.

In Fig. 1*a* we show a composite picture of an extended line-by-line⁸ spectrum obtained through the large aperture (20 × 10 arc s oval shape) of IUE, in the range 1,200–1,500 Å. Figure 1*c* shows the three stars in the same orientation as they were observed with IUE after the supernova had faded in the far ultraviolet. Figure 1*a* is a composite of three individual spectra (see below) taken in early March. The reason we use such early spectra is the fact that the position angle of the large aperture varies at the position of the supernova by ~+1° per day so that the best spatial resolution could be achieved only at those early epochs. Comparison with later spectra shows that the contribution of the supernova at these wavelengths was ~5% or less. Figure 1 shows that the spectrum after the disappearance of the SN1987A in the far ultraviolet actually consists of two-point source spectra. Figure 1*b* shows the intensity tracing across these two spectra and the deconvolution with the skewed gaussian point spread function (PSF) of IUE and confirms our preliminary results. From Fig. 1 it is clear that the suggestion that star 1 be still present is not consistent with the optical characteristics of these stars (ref. 3; N. R. Walborn *et al.* preprint, 1987). In fact, if the upper spectrum is identified with star 1, one would have the brighter star in the ultraviolet corresponding with the fainter star in the optical. But star 1 is about 3 mag brighter than star 2 in the visual. Therefore, barring completely unacceptable behaviour to connect the optical to the ultraviolet spectrum, one might try to explain such a result in terms of an especially high extinction in front of star 1. This would require a very large spatial variation in the interstellar extinction, which finds no support in the optical photometry nor in the shape of the ultraviolet spectra. These arguments led to the preliminary conclusion that star 1 was probably the supernova progenitor^{6,7}.

The results which are discussed below will put these suggestions on a fully quantitative basis. The fact that the two spectra can be separated (the pixel size perpendicular to the dispersion in the IUE spectra is 1.078 arc s), points to a separation larger than 3 arc s, which is the separation one should expect if the observed ultraviolet spectra were those of stars 1 and 2. We have used three images obtained with the short wavelength prime (SWP) camera (30422, 30427, 30472 taken on 2, 3 and 9 March 1987, respectively) to determine the separation between the two spectra. We have performed a sophisticated deconvolution, using the skewed PSF of the SWP camera⁹ in the broad-

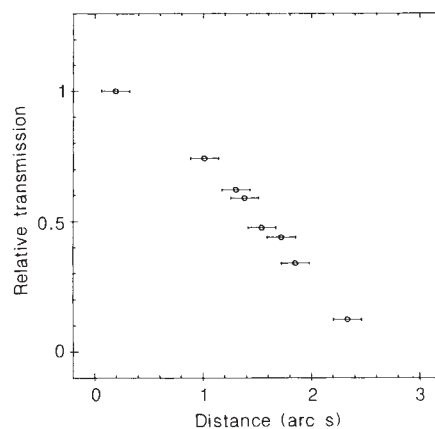


Fig. 2 The relative transmission of the IUE short wavelength small aperture as a function of the distance from the centre. The error bars indicate the measured uncertainty in the pointing of IUE.

band region 1,300–1,500 Å where the spatial resolution of the SWP camera is best. We obtained a value of 4.5 ± 0.2 arc s for the separation, and 1.75 ± 0.12 for the mean intensity ratio between the two stars. These results are to be compared with the optical determinations of the separations (in the projection perpendicular to the dispersion line) of 3 arc s between stars 1 and 2, and 4.4 arc s between stars 2 and 3; and with an optical flux ratio of 0.06 for star 2 to star 1 and of 1.4 for star 2 to star 3. All of this strongly suggests that in Fig. 1 we see the spectra of stars 2 and 3. Still there could be some doubt due to the known 5–10% uncertainty in the spatial scale of IUE⁸ and because no direct comparison can be made at present between the field in the optical and in the ultraviolet. But by using the optical image of the supernova to position the IUE spacecraft, we can use the small aperture in the IUE focal plate (3 arc s) to make flux measurements which can be directly compared with the optical results available. In this way it is possible to analyse this problem from the flux point of view. For this purpose on 16 April we have made an IUE observation with the small aperture (sap) centred on the position of star 1, using SN1987A in the optical to define the direction in which the spacecraft is pointing. Extreme care has been taken to stabilize the pointing of IUE and all conditions were optimally maintained for the full 320-min exposure. We obtained an average (1,300–1,500 Å) flux of $F(\text{sap}) = 0.50 \pm 0.05 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Å}^{-1}$. Moreover, the transmission of the small aperture as a function of the distance from the centre was measured separately on a standard star (BD + 284211) under nearly identical conditions along the line connecting stars 1 and 3. The results are shown in Fig. 2. The fluxes measured with the large and the small apertures of IUE are governed by the following equalities:

$$F(\text{lap}) = F(1) + F(2) + F(3) \quad (1)$$

$$F(\text{sap}) = a[F(1) + bF(3)] \quad (2)$$

where $F(\text{lap}) = 5.94$ (in units of $10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Å}^{-1}$) is the average (1,300–1,500 Å) flux measured in several large aperture spectra around the epoch of the small aperture observation. The numerical indexes refer to the stars in Fig. 1*c*. The constants $a = 0.51 \pm 0.05$ and $b = 0.52 \pm 0.02$ represent the absolute point source transmission of the IUE small aperture at the centre (at the position of star 1) as determined from several observations of BD + 284211, and the relative transmission at an offset corresponding to the separation of star 3 from star 1. The flux of star 2 was measured through a deconvolution of the large aperture spectrum with the skewed IUE point spectral function (PSF) and was found to be $F(2) = 3.78$ (same units) in the interval 1,300–1,500 Å. This value is in excellent agreement with the

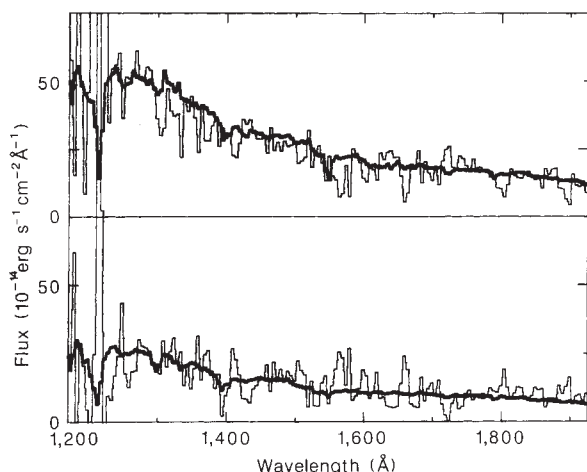


Fig. 3 The dereddened spectra of stars 2 (top) and 3 (bottom). The thick lines are spectra of standard stars, scaled to the optical magnitudes of stars 2 and 3. The spectral types are B0V for star 2 and B1.5V for star 3.

direct determination⁵ made by using the small aperture centred on star 2 (4.5, same units, at 1,300Å). In this way the two equations are based exclusively on measured quantities and can be solved reliably for $F(1)$ and $F(3)$, while the independently measured flux of star 2 acts as a check on the results. Therefore, by solving equations (1) and (2) we can determine the actual ultraviolet fluxes (1,300–1,500Å) of stars 1 and 3 after the supernova explosion. We obtain:

$$F(1) = -0.3 \pm 0.4$$

$$F(3) = 2.5 \pm 0.4$$

where the errors were computed as

$$\sigma_F^2 = \sum_i \sigma_{x_i}^2 \times (\partial F / \partial x_i)^2$$

and take into account the 10% flux uncertainty in IUE spectra, the small aperture transmission errors and a 0.10 arc s uncertainty in the measure of the distance between stars 1 and 3—this last uncertainty was included in the error on b (± 0.05). It appears that the flux of star 1 is effectively zero and that star 3 can account for the whole flux observed in the small aperture. Considering that the expected flux for a B3I star of $V=12.3$ with an assumed $E(B-V)=0.20$ (ref. 6) is $F(1)_{\text{predicted}} = 7.5 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Å}^{-1}$, then it is quite clear that star 1 is not detected in the small aperture spectrum and that there can be no question about the fact that it has disappeared.

As a final test of these results, we have isolated the spectra of stars 2 and 3 by scaling the small aperture spectrum to the value obtained for star 3 through the deconvolution, and then subtracting it from the average large aperture spectrum to obtain star 2. Although the spectral features are dominated by the poor signal to noise in the small aperture spectrum, the resulting spectra (see Fig. 3) should give a good overall representation of the two stars. The spectra were dereddened using⁶ $E(B-V)=0.05$ for the galaxy and $E(B-V)=0.15$ for the LMC (30 Doradus extinction law¹⁰). In Fig 3, the thick line is a spectrum of a standard star from the IUE low dispersion spectra reference atlas¹¹: HD36512 (B0V) for star 2 and HD74273 (B1.5V) for star 3. These stars were selected on the basis of their dereddened colours and absolute magnitudes best matching those of stars 2 and 3 (for a distance modulus to the LMC of 18.6). Their IUE spectra were dereddened for the appropriate $E(B-V)$ ¹² and scaled to the dereddened V magnitude of stars 2 and 3. The agreement is quite satisfactory. Therefore, we feel confident in attributing a spectral classification of B0V and B1.5V to star 2 and star 3, respectively. It is rewarding to realize that the spectral type for star 2 is the same as derived for this star on the basis of optical photometry³.

We conclude that the separation, the fluxes and the spectral appearance of the two sources observed with IUE when the SN1987A had faded demonstrate that star 2 and star 3 are the only ones still present in the field. This fully confirms our earlier suggestion⁶ that the B3I star (star 1) in Sk-69 202 has disappeared after the explosion of SN1987A because it is actually its progenitor.

Received 16 June; accepted 7 July 1987.

- McNaught, R. H. *IAU Circ. No. 4316* (1987).
- McNaught, R. H. & Morel, M. *IAU Circ. No. 4317* (1987).
- West, R. M. *et al. Astr. Astrophys. 177*, L1-L3 (1987).
- Gry, C. *et al. IAU Circ. No. 4327* (1987).
- Sonneborn, G. & Kirshner, R. *IAU Circ. No. 4333* (1987).
- Panagia, N. *et al. Astr. Astrophys. 177*, L25-L28 (1987).
- Kirshner, R. P., Sonneborn, G., Crenshaw, D. M. & Nassiopoulos, G. E. *Astrophys. J. Lett.* (in the press).
- Munoz, J.R. in *IUE Image Processing Manual* European edn, 11/1-11/21 (ed. Clavel, J.) (Madrid, 1987).
- Cassatella, A. *et al. Astr. Astrophys. 144*, 335-342.
- Koorneef, J. & Code, A. D. *Astrophys. J. 247*, 860-868 (1981).
- Heck, A. *et al. IUE Low-Dispersion Spectra Reference Atlas* ESA SP-1052 (ESA, Paris, 1984).
- Wu, C.-C. *NASA IUE Newsletter 22*, 1-324 (1983).

Light-curve models for supernova SN1987A in the Large Magellanic Cloud

T. Shigeyama, K. Nomoto, M. Hashimoto & D. Sugimoto

Department of Earth Science and Astronomy,
College of Arts and Sciences, University of Tokyo, Meguro-ku,
Tokyo 153, Japan

Neutrino bursts from the supernova SN1987A have been reported¹⁻³ which provide qualitatively new information as to the time interval between the core collapse and the optical flaring of the supernova up to about 6 mag⁴. The interval that we take as three hours^{1,2} imposes conditions on the radius of the progenitor star and the energetics of the explosion. As for the brightness and its time change, the relatively less bright visual magnitude at the plateau in the light curve during the initial eight days or so (refs 5, 29) as well. The subsequent light curve after eight days is quite unique; the bolometric luminosity increases until ~65 days (ref. 29) but thereafter it seems to level off gradually (J. Walsh and R. Stathakis, personal communication; M.A. Dopita, personal communication). To maintain the bolometric luminosity of the star means that certain conditions should be imposed on the energy source. By computing the propagation of shock wave, the subsequent expansion of ejected matter and the optical light curve, we find that the early light curve can be accounted for by the diffusive release of energy deposited by the shock wave, that the progenitor's radius should be as small as $\sim (1-3) \times 10^{12} \text{ cm}$, and that the explosion energy per unit mass should be relatively large, $\sim (2-3) \times 10^{51} \text{ erg}$ for the ejected matter of 7–10 $M_{\odot}$. The light curve after eight days is well reproduced by using a model with constant energy input. The energy input may be due to the activity of an embedded pulsar. The light curves calculated by invoking the radioactive decay of ⁵⁶Ni and ⁵⁶Co are less satisfactory in accounting for the nearly flat portion observed after 65 days. Supposing Sk-69 202 is the progenitor of SN1987A^{13,14}, its explosion energy should be $> 2 \times 10^{51} \text{ erg}$; this suggests that a prompt shock mechanism forms a neutron star.

We set the time $t=0$ when the shock wave is generated in the deep interior of the stellar core. The time when the shock arrives at the original surface of the star is denoted by t_{prop} . To determine t_{prop} we carried out hydrodynamic calculations on different models of progenitor stars, that is, for different values of the initial radius R_0 and the ejected mass M , and for different

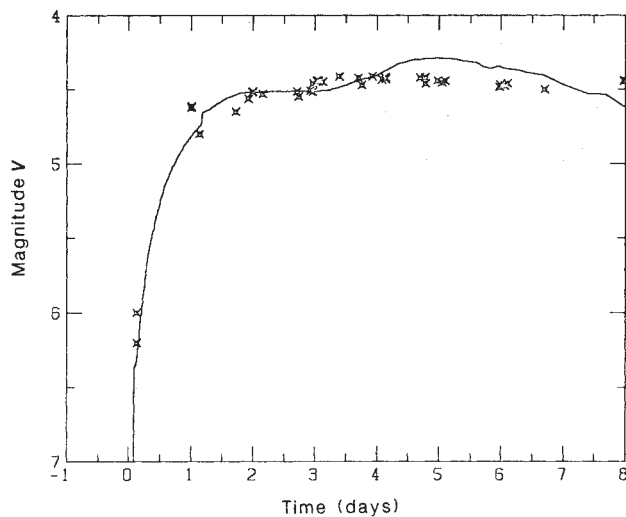


Fig. 1 Change in the calculated visual magnitude for the standard model with $M = 7 M_{\odot}$, $R = 1.5 \times 10^{12}$ cm, and $E = 2.5 \times 10^{51}$ erg (solid line)¹⁹. Star marks indicate the observations at the European Southern Observatory⁵, Cerro Tololo Inter-American Observatory¹⁹ and South African Astronomical Observatory²⁹ except for the two points at $t = 3$ h on films.

energies of explosion E . The energy E was deposited instantaneously in the central region of the polytrope of index 3 to generate a strong shock wave. The envelope structure of the actual presupernova star with a radius of moderate size is well approximated by this model⁶. The propagation of the shock wave was calculated with a hydrodynamic code⁷ with a modification to include flux-limited diffusion⁸ until the shock wave arrives close to the original photosphere. Then the value of t_{prop} was determined, which can be approximated by

$$t_{\text{prop}} \sim 35 \text{ min } (R_0/1.5 \times 10^{12} \text{ cm}) \times [(M/7 M_{\odot})/(E/2.5 \times 10^{51} \text{ erg})]^{1/2} \quad (1)$$

A progenitor radius $> 10^{13}$ cm is ruled out simply using the condition $t_{\text{prop}} < 3$ h (Kamiokande time will be used hereafter).

When the shock wave arrives at the photosphere, the temperature behind the shock is as high as 10^6 K so that most of the radiation is emitted from the extreme ultraviolet to the soft X-ray band. The optical flare-up of the supernova can be seen only after the ejected gas and the radiation field have expanded so that the temperature becomes lower and the radius of the photosphere becomes larger. Therefore, we need to calculate the time from t_{prop} to the optical flare-up.

During the expansion stage, the colour photosphere is determined for each wavelength by the condition

$$\int_r^{\infty} [3 \kappa_{\nu} (\kappa_{\nu} + \kappa_{\text{es}})]^{1/2} dr = 1 \quad (2)$$

where κ_{es} and κ_{ν} denote the electron scattering opacity and the absorption opacity at the frequency ν , respectively⁹⁻¹¹. This colour photosphere determines the colour temperature T_{colour} of the emitted photon. On the other hand, the emitted flux is determined by the temperature thereof and the diffusion velocity of the photon c/τ_{es} with τ_{es} being the electron scattering optical depth. The total flux integrated over all frequencies is practically equal to the diffusion luminosity which is calculated at the scattering photosphere, that is, at $\tau_{\text{es}} \sim 1$, because the energy flux due to the diffusion of the photon is approximately constant through the expanding envelope^{10,11}. The advection luminosity, the flux of the radiation field which is trapped within a mass element of the gas and transported with the mass motion of the gas, amounts to $4v/c$ times the diffusion luminosity. This advection luminosity is calculated at the scattering photosphere and added to the diffusion luminosity to obtain total luminosity^{10,11}.

Details of the hydrodynamic calculation will be discussed

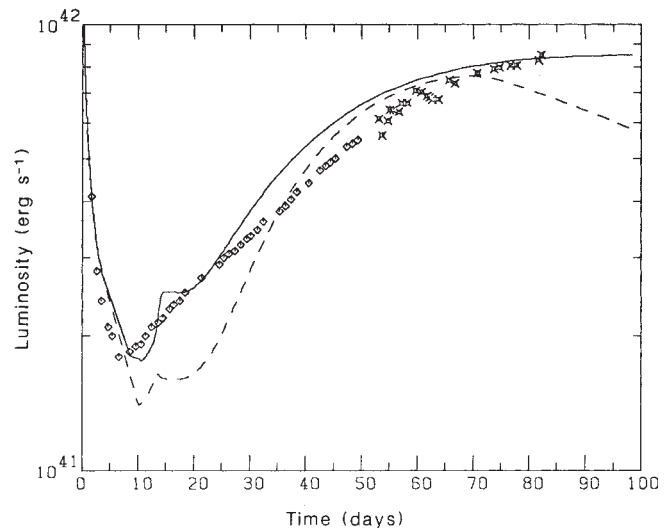


Fig. 2 Change in the calculated bolometric luminosity for the standard model with constant energy input ($L_{\text{md}} = 8.5 \times 10^{41}$ erg s⁻¹) due to pulsar activity (solid) and with radioactive decays of ^{56}Ni and ^{56}Co (dashed). For the latter model with radioactive decay, the mass of ^{56}Ni is $0.1 M_{\odot}$ and the opacity of the helium layer is enhanced to $0.35 \text{ cm}^2 \text{ g}^{-1}$. Observational data are taken from South African Astronomical Observatory²⁹ (squares) and Anglo-Australian Observatory (J. Walsh and R. Stathakis, personal communication) and Mt Stromlo and Siding Spring (personal communication) denoted by star marks.

elsewhere. Here we discuss mainly a standard model which is obtained for $R_0 = 1.5 \times 10^{12}$ cm, $E = 2.5 \times 10^{51}$ erg and $M = 7 M_{\odot}$. We assume that the ejecta consist of an innermost heavy-element layer of $2.4 M_{\odot}$, a helium-rich layer of $2.2 M_{\odot}$, and a hydrogen-rich envelope of $2.4 M_{\odot}$ ($X = 0.74$ and $Z = 0.001$, where X and Z denote the concentrations of hydrogen and heavy elements, respectively). For the heavy-element layer a constant opacity of $0.25 \text{ cm}^2 \text{ g}^{-1}$ is assumed. For the helium layer, the helium-line opacity of $\kappa = 0.009 \text{ cm}^2 \text{ g}^{-1}$ sets the lower bound value of the opacity (Los Alamos opacity table for Nomoto mixture¹²). These line opacities are highly uncertain because of the variation in Doppler shift for a large velocity gradient. The chemical structure is taken from the presupernova model of a $6 M_{\odot}$ helium star⁶; it is designed to simulate a star of $M_{\text{ms}} \sim 20 M_{\odot}$ on its main sequence which subsequently loses a significant fraction of its hydrogen-rich envelope and finally leaves a neutron star of $\sim 1.4 M_{\odot}$ after the explosion. The luminosity of Sk-69 202, corresponds to such a mass for the progenitor star (refs 15-17; E. K. Grassberg *et al.*, preprint, 1987).

For this standard model, the shock wave established a radiation field with an energy of 1.2×10^{51} erg. Figure 1 shows the time change in the calculated V magnitude where a distance modulus of 18.5 and an extinction of 0.6 are adopted¹⁸. The star marks indicate the observations at the European Southern Observatory⁵, Cerro Tololo Inter-American Observatory¹⁹ and South African Astronomical Observatory²⁹ except for the two points measured on film at $t = 3$ h⁴. Our theoretical light curve is in good agreement with the observations despite the approximate presupernova configuration adopted in the present calculation.

After the shock wave heats the stellar surface, the theoretical visual luminosity quickly rises and reaches a plateau in one day. The relatively high velocity of the expansion due to large E/M explains the rapid decline of the colour temperature in early phase, especially at $t = 3$ h. During the plateau at about $t = 1-8$ days, the visual luminosity stays almost constant because the photosphere becomes deeper in lagrangian coordinates as the recombination front propagates inwards, which compensates for the effect of expansion. In contrast to the explosion of a red

supergiant which is usually considered for normal type II supernovae, the radiation field was initially set up in a small volume and later suffered from a decrease in temperature due to an adiabatic expansion. This explains the relatively low luminous plateau. In this way, the early light curve up to $t \sim 8$ days can be accounted for by the diffusive release of energy which was deposited in the hydrogen-rich envelope during the passage of the shock wave.

For other parameters, reasonable agreement between the theoretical early light curve and observations is obtained for $R_0 = (1-3) \times 10^{12}$ cm and $E = (2-3) \times 10^{51}$ erg for $M = 7-10 M_\odot$.

After these early stages, the observations show the increase in the bolometric luminosity as summarized in Fig. 2 (ref. 29; J. Walsh, R. Stathakis and M. A. Dopita, personal communications). If no additional heating is included in our model, the luminosity starts to decrease at about $t \sim 10$ d. This implies that there should exist an energy source that continually heats the expanding star. As a possible energy source, we consider two cases: (1) a newly formed pulsar^{20,21} and (2) the radioactive decays of ^{56}Ni and ^{56}Co . The magnetic dipole radiation from the pulsar is approximated to deposit energy in the innermost layers at a constant rate ($L_{\text{md}} = 8.5 \times 10^{41}$ erg s⁻¹) as its decay timescale is estimated as a few years²⁰. The radioactive decay and the deposition of γ rays are treated by applying the approximate deposition function²².

Figure 2 compares the calculated bolometric light curve with observations (ref. 29; J. Walsh, R. Stathakis and M. A. Dopita, personal communications). The photosphere contracts as the recombination front proceeds inward through the hydrogen-rich envelope. When the recombination front and thus the photosphere come close to the helium layer, the theoretical bolometric luminosity starts to increase. The photosphere reaches the helium layer at $t = 15$ d and then the heavy-element layer at $t = 40$ d. The bolometric luminosity increases until it approaches the peak value of $L = L_{\text{md}}$. It is in reasonable agreement with observations, especially near the peak. (Better agreement is obtained if a larger line opacity of $\kappa = 0.05$ cm² g⁻¹ is assumed for the helium layer.) For a model with a hydrogen-rich envelope of $5.4 M_\odot$, that is, for the ejected mass of $M = 10 M_\odot$, the agreement is less satisfactory because the photosphere reaches the helium layer at a later stage. This suggests that a significant fraction of the hydrogen-rich envelope should have been ejected before the presupernova stage.

If indeed the later luminosity increase is due to the pulsar activity, our power of L_{md} for SN1987A is appreciably lower than $\sim 10^{44}$ erg s⁻¹ adopted by Bodenheimer and Ostriker²⁰. This implies that the magnetic field and/or the initial angular velocity should be smaller. Our model predicts that this second plateau in luminosity will continue until the expanding gas becomes thin and the thermal dissipation of the magnetic dipole radiation starts to decrease³⁰. At around that time, the optical luminosity will decrease while the non-thermal ultraviolet and X-ray luminosity will increase^{20,23}. Moreover, very-high-energy γ -rays produced from the interaction between high-energy particles and the ejected gas will be detected²⁴.

The theoretical light curve with radioactive decay of ^{56}Ni and ^{56}Co was also calculated assuming the production of $0.10 M_\odot$ ^{56}Ni . If the Los Alamos floor value of 0.009 cm² g⁻¹ is adopted for the helium-line opacity, a peak luminosity of 9×10^{41} erg s⁻¹ is reached as early as about $t = 50$ d. To improve it, we may consider the case where the opacity in the helium layer is as high as $\kappa = 0.35$ cm² g⁻¹. This might be possible if we include the effects of a velocity gradient on the line opacity. Moreover, a slower velocity helps it because the shape of the light curve depends on $\kappa M/v$ where v denotes the characteristic expansion velocity²⁵. Such a slower expansion velocity of the core may be expected if we consider more realistic models where the density of the helium layer is significantly higher than in our model²⁶. The result for such an enhanced opacity is shown by the dashed line in Fig. 2. Even for this case the agreement with observations

is poor. The increase in luminosity is too steep, and, in particular the decline, which is due to the decline in the γ -ray deposition, is too fast.

Our results imply that the pulsar model is able to account for observations, in particular the plateau-like peak, better than the radioactive-decay model. In addition, it suggests that the amount of ^{56}Ni is smaller than $\sim 0.1 M_\odot$. It is still very likely that some ^{56}Ni is produced in the massive star model so that the X rays and γ rays that originated from any radioactive decay will be detected²³ as well as those originating from pulsar activity. If we use a more realistic presupernova model with denser cores, it is possible to obtain better agreement between the radioactive model and the observations than in the present study (T.S. and K.N., in preparation).

We consider the evolutionary implications of our light-curve model. Suppose that the progenitor is Sk-69 202 and thus as massive as $M_{\text{ms}} \sim 15-20 M_\odot$. Then the explosion energy must be larger than 2×10^{51} erg. This suggests that the mechanism that transforms collapse into explosion may be a prompt shock mechanism rather than late-time neutrino heating because the latter mechanism produces at most $\sim 1 \times 10^{51}$ erg even including convection in neutrino transport (J. R. Willson and R. W. Mayle, preprint, 1987). On the other hand, a prompt explosion with energy of 2.2×10^{51} erg is obtained (E. Baron *et al.*, preprint, 1987) from the collapse of $1.2 M_\odot$ iron core⁶. Such a small iron core was obtained from improvements of equation of state, electron capture and silicon burning reaction network⁶.

The evolutionary process leading to a progenitor radius of $R_0 \sim 10^{12}$ cm is not yet quite clear and needs further study. In the Large Magellanic Cloud, there exist many bright red supergiants corresponding to stars up to $M_{\text{ms}} \sim 50 M_\odot$ ²⁷. Therefore, unless Sk-69 202 is a close binary, it is likely that the progenitor evolved first to a red supergiant stage and, after losing a substantial fraction of its hydrogen-rich envelope, returned to a blue supergiant (D. J. Faulkner and P. R. Wood, preprint, 1987). The early onset of brightening ($t \sim 8$ d) is consistent with the relatively small hydrogen mass. The low metallicity would also play a role in determining R_0 (refs 16, 17 and 26).

The above model predicts the existence of low-density inner circumstellar matter and denser outer matter around SN1987A²⁸. The latter originated from a slow wind during its preceding red supergiant stage (or Roche lobe overflow if it is a close binary) which was caught up with by the higher-velocity wind during the blue supergiant stages. Therefore, if we wait until the shock reaches the outer circumstellar matter, X rays will be emitted and detections of different elements may become possible²⁸.

To summarize, our massive star model for SN1987A suggests: (1) the small radius of the progenitor, (2) a relatively large explosion energy per unit mass, (3) extensive mass loss before the explosion, and (4) the early onset of pulsar activity. We expect the detection of X rays and γ rays from the new-born pulsar, radioactivity and interaction between ejecta and outer circumstellar shell.

We are very grateful to Drs Jeremy Walsh, Raylee Stathakis, and Mike Dopita for sending us observational data obtained at the Anglo-Australian Observatory and Mt Stromlo and Siding Spring Observatory. We thank Professors G. E. Brown and J. Ostriker for stimulating discussions and Shiori Kumagi and Hisashi Shiba for their assistance to our numerical calculation. Thanks also to Drs S. Miyaji, T. Ebisuzaki, F.-K. Thieleman, P. Sutherland, and T. Hanawa for useful correspondence. This work has been supported in part by the Japanese Ministry of Education, Science and Culture and by the Space Data Analysis Center, Institute of Space and Astronautical Sciences.

Received 9 April; accepted 8 June 1987.

- Hirata, K. *et al.* (Kamiokande-II collaboration) *Phys. Rev. Lett.* **58**, 1490-1493 (1987).
- Bionta *et al.* *Phys. Rev. Lett.* **58**, 1494-1496 (1987).
- Castagnoli, C. *et al.* *IAU Circ. No.* 4323 (1987).
- McNaught, R. H. *IAU Circ. No.* 4316 (1987).
- Cristiani, S. *et al.* *Astr. Astrophys.* **177**, L5-L8 (1987).

6. Nomoto, K. & Hashimoto, M. *Prog. Part. Nucl. Phys.* **17**, 267–285 (1986); *Proc. Japan–France Seminar on Chemical Evolution of Galaxies with Active Star Formation*, Sci. Rep. Tohoku Univ. 7 (ed. Takakubo, K.) 259–276 (Tohoku Univ., 1987); *Proc. Beihai Conf. Supernovae*, Physics Report (ed. Brown, G. E.) (North-Holland, in the press).
7. Yahil, A., Johnston, M. D. & Burrows, A. *J. comput. Phys.* (in the press).
8. Pomraning, G. C. *Astrophys. J.* **266**, 841–847 (1983).
9. Hershkovitz, S., Linder, E. & Wagoner, R. V. *Astrophys. J.* **303**, 800–809 (1986).
10. Ebisuzaki, T., Hanawa, T. & Sugimoto, D. *Publ. astr. Soc. Jap.* **35**, 551–566 (1984).
11. Ebisuzaki, T. & Nomoto, K. *Astrophys. J.* **305**, L67–70 (1986).
12. Huebner, W., Merts, A., Magee, M. Jr & Argo, M. *Los Alamos Sci. Lab. Rep.* No. LA6760M (1977).
13. Panagia, N. *et al. Astr. Astrophys.* **177**, L25–L28 (1987).
14. Kirshner, R., Sonneborn, G., Crenshaw, D. M. & Nassiopoulos, G. E. *Astrophys. J.* (in the press).
15. Woosley, S. E., Pinto, P. A., Martin, P. G. & Weaver, T. A. *Astrophys. J.* (in the press).
16. Hillebrandt, W., Hoflich, P., Truran, J. W. & Weiss, A. *Nature* **327**, 597–600 (1987). (in the press).
17. Arnett, W. D. *Astrophys. J.* (submitted).
18. West, R. M., Lauberts, A., Jorgensen, H. E. & Schuster, H. E. *Astr. Astrophys.* **177**, L1–L3 (1987).
19. Suntzeff, N. *et al. IAU Circ.* No. 4328 (1987).
20. Bodenheimer, P. & Ostriker, J. P. *Astrophys. J.* **181**, 465–471 (1974).
21. Ostriker, J. P. *Nature* **327**, 287–288 (1987).
22. Weaver, T. A., Axelrod, T. A. & Woosley, S. E. *Type I Supernovae* (ed. Wheeler, J. C.) 113–154 (University of Texas, Austin, 1987).
23. McCray, R., Shull, J. M. & Sutherland, P. *Astrophys. J.* **317**, L73–L77 (1987).
24. Sato, H. *Prog. theor. Phys.* (submitted).
25. Arnett, W. D. *Astrophys. J.* **253**, 785–797 (1982).
26. Woosley, S. E., Pinto, P. A. & Ensmann, L. *Astrophys. J.* (submitted).
27. Humphreys, R. M. & Davidson, K. *Astrophys. J.* **232**, 409–420 (1987).
28. Itoh, H., Hayakawa, S., Masai, K. & Nomoto, K. *Publ. Astr. Soc. J.* **39**, No. 3 (1987).
29. Menzies, J. M. *et al. Astr. Astrophys.* (in the press).
30. Gaffet, B. *Astrophys. J.* **216**, 852 (1978).

An interacting binary model for SN1987A

A. C. Fabian, M. J. Rees, E. P. J. Van den Heuvel* & J. Van Paradijs*

Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK
*Astronomical Institute, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands

The recent supernova¹ in the Large Magellanic Cloud (LMC), SN1987A, is positionally coincident^{2–4} with the B3 I supergiant star, Sk–69 202. Ultraviolet observations⁵ of this region made with the International Ultraviolet Explorer satellite after the explosion now show this star to have disappeared. Initial doubts that a blue supergiant could be a supernova progenitor may be overcome by calculations^{6–8} of the evolution of massive stars at reduced metallicity, as expected in the Magellanic Clouds. These models also account for the relatively low luminosity of the supernova. Here we present an alternative model for SN1987A in which the progenitor was a close binary companion of Sk–69 202.

It should be emphasized that single star models still have considerable uncertainties due to problems in modelling convection and opacities for example, so that there is as yet no agreement as to whether the blue supergiant at the supernova stage had always been a blue supergiant^{6,7} or has returned from the red supergiant phase⁸. A binary progenitor has been previously discussed briefly in refs 4 and 8 and by M. Dopita (personal communication). It is well known that a significant fraction of massive stars are in close binary systems. Several definite predictions follow if the progenitor of SN1987A was a close binary companion of Sk–69 202. First, the blue supergiant should reappear within the next year or so as the supernova photosphere shrinks. Second, if, as we expect, the binary system has remained bound, the supergiant should be found in an eccentric orbit and will evolve into a massive X-ray binary, similar to 4U1223–62 (which has a B1.5 Ia supergiant companion) or 1E1145.1–6141 (which has a B2 Ia supergiant companion, ref. 9).

The expected formation process of massive X-ray binaries has been discussed and reviewed elsewhere¹⁰. Mass transfer from the more massive, and thus more evolved, member of a close binary system consisting of O–B stars with a mass in the range of 15–30 $M_{\odot}$ leads to it being reduced to a helium star of 4–10 $M_{\odot}$. The system may appear as a Wolf–Rayet binary *en route* to this point, although extensive wind-driven mass loss may terminate well before the supernova explosion if little stellar

envelope remains on the helium star. Also, if the mass of the helium star is $\leq 6 M_{\odot}$, it may never appear as a Wolf–Rayet star, as no Wolf–Rayet stars with masses $< 6 M_{\odot}$ are known¹¹. Nevertheless, such helium stars must certainly exist, as they are produced in a straightforward way by close binary mass transfer. A helium star with a mass $< 6 M_{\odot}$ could dominate the extreme ultraviolet output of the binary system but not the visible spectrum owing to its small size. It is possible that the orbital period of the progenitor system is accessible from a careful study of archival plates (a study of 40 European Southern Observatory (ESO) Schmidt plates carried out so far² revealed no variability with an amplitude > 0.5 mag).

The supernova outburst occurs when the core of this star collapses. The mass of the ejected supernova shell will be small, but enhanced by matter picked up from the (now) more massive companion, identified as Sk–69 202. We estimate that the amount of mass stripped by the impact of the supernova shell, by means of the formalism given in ref. 12 (and justified by three-dimensional hydrodynamical calculations¹²). We find that for a mean velocity of the supernova shell of 1.4×10^4 km s^{–1}, an ejected shell mass of 4.5 $M_{\odot}$, and a separation between the components of twice the radius of the supergiant, $\sim 1 M_{\odot}$ of hydrogen-rich matter will be stripped from the envelope of a 15–20 $M_{\odot}$ companion. An additional $\sim 0.3 M_{\odot}$ may be lost by ablation (taking the corrections given in ref. 13 into account). The envelope of this star is an obvious source for the hydrogen observed in the spectrum of SN1987A. The reduced luminosity of the supernova is presumed to be due to the small size of the helium star. We require that much of the shock energy deposited in the stellar envelope was lost in adiabatic expansion and that the mass of ⁵⁶Ni ejected was small (see ref. 14 for a brief discussion of dim supernovae). The matter stripped from the companion will further modify the light curve. A prediction of our model is that the supernova ejecta contain roughly equal numbers of hydrogen and helium atoms, although the hydrogen may be concentrated preferentially along the line connecting the components at the time of the explosion.

The supernova explosion of the helium star will have increased the orbital period and eccentricity of the binary system and given it a runaway velocity. All of these properties should be optically observable after the photosphere of the supernova has shrunk back within the binary system so that Sk–69 202 is again visible. The photosphere was at radii of 10^{15} cm about 7 months after the outburst. The Thomson depth of a uniform sphere of mass $m M_{\odot}$ expanding at 10^9 cm s^{–1} reduces below unity in $4 m^{0.5}$ months. As the ultraviolet photospheric opacity of the expanding remnant will be higher than this^{15,16} and the density non-uniform, this is a minimum timescale on which Sk–69 202 can reappear. About one year may be a reasonable estimate. It will take a similar time for the remnant luminosity to drop sufficiently for Sk–69 202 to be detected at longer wavelengths. The diffuse supernova remnant should be asymmetrical due to interaction with the binary companion. There is already tentative evidence for this from optical polarization measurements¹⁷ and speckle interferometry¹⁸. The shapes of the emission lines may also be a useful diagnostic in this respect.

A neutron star, or less probably a black hole, is expected to remain from the explosion (K. Sato and H. Suzuki, preprint, 1987; A. Burrows and J. M. Lattimer, Steward Observatory, 1987). A black hole may immediately begin to accrete matter from the companion, probably in an orbitally periodic manner. A neutron star is likely to be rapidly spinning and magnetic and therefore potentially observable as a pulsar. Radio emission is probably quenched in the wind from Sk–69 202 and ‘propeller’ spindown occurs¹⁹. The best opportunity for observing it seems to be as a rotation-powered Crab-like X-ray pulsar such as the young LMC pulsar PSR0540–693. If this hypothesis is correct then we will have an unprecedented opportunity to observe this important phase in the life of an X-ray binary. A combination of optical spectroscopy and possibly X-ray measurements may

allow the orbital position of the supernova progenitor at the time of the explosion to be estimated if tidal effects are ignored. This should allow the mass loss, ablation and rocket effects of Sk-69 202 to be accurately determined, especially if the pre-supernova orbital parameters can be determined as outlined above. (The ablation effects can be derived from its change in radius with respect to the presupernova stage.) Finally, we note that if Sk-69 202 emerges again and is found to be in orbit then we will be able to measure the mass of the collapsed object.

Received 13 May, accepted 2 July 1987.

- Madare, B. & Kunkel, W. *IAU Circ.* No. 4316 (1987).
- West, R. M., Lauberts, A., Jorgensen, H. E. & Schuster, H.-E. *Astr. Astrophys.* (submitted).
- White, G. L. & Malin, D. F. *IAU Circ.* No. 4330 (1987).
- White, G. L. & Malin, D. F. *Nature* **327**, 36-38 (1987).
- Kirshner, R. P., Nassiopoulos, G. E., Sonneborn, G. & Crenshaw, D. M. *Astrophys. J.* (in the press).
- Arnett, D. A. *Astrophys. J.* (in the press).
- Hillebrandt, W., Höflich, P., Truran, J. W. & Weiss, A. *Nature* (in the press).
- Woosley, S. E., Pinto, P. A. & Ensmann, L. *Astrophys. J.* (submitted).
- Van Paradijs, J. A. in *Accretion-Driven Stellar X-ray Sources* (eds Lewin, W. H. G. & van den Heuvel, E. P. J.) 189-260 (Cambridge University Press, 1983).
- van den Heuvel, E. P. J. in *Accretion-Driven Stellar X-ray Sources* (eds Lewin, W. H. G. & van den Heuvel, E. P. J.) 303-341 (Cambridge University Press, 1983).
- Massey, P. *Astrophys. J.* **246**, 153-160 (1981).
- Wheeler, J. C., Lecar, M. & McKee, C. F. *Astrophys. J.* **200**, 145 (1975).
- Fryxell, B. A. & Arnett, W. D. *Astrophys. J.* **243**, 994-1002 (1981).
- Arnett, D. A. in *Supernovae, A Survey of Current Research* (eds Rees, M. J. & Stoneham, R. J.) 235-252 (Reidel, Dordrecht 1982).
- Karp, A. H., Lasher, G., Chan, K. L. & Salpeter, E. E. *Astrophys. J.* **214**, 161-178 (1977).
- McCray, R., Shull, J. M. & Sutherland, P. *Astrophys. J.* **317**, L73-L77 (1987).
- Schwarz, H. E. & Mundt, R. *Astr. Astrophys.* (submitted).
- Karovska, M., Nisenson, P., Noyes, R. & Papaliolios, C. *IAU Circ.* No. 4382 (1987).
- Illarionov, A. F. & Sunyaev, R. A. *Astr. Astrophys.* **39**, 185-195 (1975).

Radio caustics from localized interstellar medium plasma structures

Roger W. Romani*, Roger D. Blandford* & James M. Cordes†

* Theoretical Astrophysics 130-33, California Institute of Technology, Pasadena, California 91125, USA

† Astronomy Department and NAIC, Cornell University, Ithaca, New York 14853, USA

In a study of 36 extragalactic radio sources observed over 7 years with the Green Bank Interferometer at 2.7 GHz and 8.1 GHz, Fiedler *et al.*¹ have detected unusual variations in the light curves of several sources. The most dramatic event, found in the flux record of the quasar 0954+658, showed large modulations at both frequencies. At 8.1 GHz, there are four spikes over a period ~80 days in each of which the intensity increases roughly threefold. At 2.7 GHz, two broad maxima of width ~30 days bracket the high-frequency event, with a total duration of ~120 days. At least two other sources show similar 2.7-GHz variations. Fiedler *et al.* argue cogently that these variations are unlikely to be intrinsic to the sources and that refractive scintillation in the interstellar medium is probably responsible. They postulate the existence of large-scale (10^{13} - 10^{14} cm) localized plasma density inhomogeneities in the interstellar medium (ISM) to explain these observations. Independent evidence for strong local fluctuations in the ISM electron density on these scales comes from the unusually large amplitude of low-frequency quasar variation², the presence of periodicities in pulsar dynamic scintillation spectra³ and non-gaussian spikes in pulsar light curves⁴. All three phenomena have been interpreted as multiple imaging and focusing by large-scale refractors. Here we present a more detailed interpretation of the event in 0954+658, propose possible sites for the refracting clouds and suggest some future observations.

The 0954+658 event can be understood in terms of geometrical optics as the signature of a strong focusing event or caustic. We picture the line of sight as passing through a

localized plasma structure having a simple gaussian electron density profile along the path of relative motion. Such a structure is a diverging radio-wave lens, causing ray crossing in the far field (analogous to the case of converging gravitational lenses postulated to explain multiple quasar imaging). The locus of points for which rays cross (in the two dimensions of time t and lens distance D) forms fourfold caustics which meet pairwise in two cusp points at a distance D_c (Fig. 1). A point source will be infinitely magnified in the geometrical optics limit for an observer located on these caustics. As the line of sight crosses each fold line there is an image pair creation/merger event and the total received flux passes through a sharp maximum. We interpret the four spikes seen at 8.1 GHz as fold crossings by an observer at a distance $D = \alpha D_c$, $\alpha \geq 3$. The characteristic ~10d rise times of these spikes are due to the convolution of the narrow caustic with a finite intrinsic source size ($\theta_{\text{int}} \sim 0.4$ mas at 5 GHz⁵). The refraction angle, and hence α , vary with frequency $\propto \nu^{-2}$ so that at 2.7 GHz the caustics separate further. The intrinsic source size should, however, also increase. (Models in which the source size is dictated by synchrotron self absorption give $\theta_{\text{int}} \propto \nu^{-1}$.) Convolution with this size at 2.7 GHz smooths over the fold pairs associated with each cusp and creates two broad ($\tau \sim 30$ day) maxima, as observed. Numerical computations with similar lens density profiles show that the general form of the light curves is quite robust for a simple clump, and that perturbations from the exact gaussian form can reproduce the observed asymmetry of the 8.1 GHz peaks. For the two other observed events, we presume that α is too small and/or θ_{int} is too large to resolve the folds at 8.1 GHz.

The plasma lenses should lie at typical distances of $D \sim 1$ kpc and the combined motion of the earth and the ISM should give a relative velocity of $100v_t$ km s⁻¹. If at an observation frequency ν , the event lasts $100\tau_2$ days and the spacing of the peaks implies $D = \alpha D_c$, then the inferred refraction angle is $\theta_{\text{ref}} \sim 6\alpha v_t \tau_2 / D_{\text{kpc}}$ mas, arising from a plasma structure of size $a \sim 10^{14} v_t \tau_2$ cm. The associated electron column density variation is $\delta N_e \sim 2\pi\theta_{\text{ref}} a \nu^2 / (r_e c^2)$. For the dramatic event in 0954+658 we infer that $\alpha \geq 3$ at 8.1 GHz and that the required column density is $\delta N_e \sim 10^{19}$ cm⁻², while for a more typical event $\alpha \sim 1$ at 2.7 GHz and $\delta N_e \sim 5 \times 10^{17}$ cm⁻². For a spherical lens the electron density is $n \sim \delta N_e / a$; less extreme densities are required if the plasma structures are elongated along the line of sight. If the elongation factor is $\eta \sim 100\eta_2$, then for a typical event the required plasma density is $n \sim 50\eta_2^{-1} \alpha (v_t \tau_2)^2 / D_{\text{kpc}}$ cm⁻³ while for 0954+658, $n \sim 10^3 \eta_2^{-1} \alpha (v_t \tau_2)^2 / D_{\text{kpc}}$ cm⁻³. As described by Fiedler *et al.* the duty cycle for these events in their data set is $f \sim 0.005$. If a caustic event is produced by a plasma filament with an aspect ratio η , aligned along the line of sight, then the cross-section for such end-on events is η^{-3} times that for filament intersections at average angles of incidence. By contrast, if the structures are locally planar, caustic events should occur η^{-2} times as often as face-on intersections. We consider the latter geometry as more probable, as it gives the greatest caustic event duty cycle for a given η and space density of the plasma lenses.

Requiring that such single high-density features occasionally dominate the distributed scattering along the line of sight contrast markedly with the conventional description of scintillation. However, it is in accord with models of the ISM incorporating high-density, cold clouds embedded in a tenuous inter-cloud medium by supernovae⁶, and we regard ionization fronts and/or cooling instabilities in old supernova remnants as reasonable sites for the lensing plasma sheets^{7,8}. This interpretation becomes more plausible if the scale height is ~1 kpc, somewhat larger than that usually inferred from observations⁹. Using the fiducial parameters for a typical plasma lens given above, the pressure in the refracting structures will be $\sim 7 \times 10^{-11}$ dyn cm⁻², for $T \sim 10^4$ K; such pressures can be found in remnants of radius $R \sim 35$ pc. As these remnants can have a covering fraction $C \sim 0.2$ at moderate galactic latitude, the probability that we are looking through an edge-on sheet in an old remnant should be

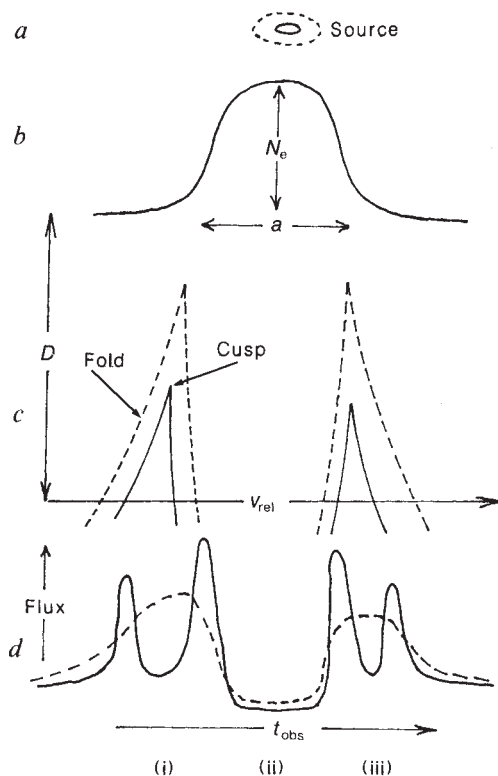


Fig. 1 Caustic formation by a localized plasma structure. A gaussian profile of excess electron density (*b*) causes ray crossing beyond $D_c \sim 1(c)$. The loci of the resultant cusp point and fold line caustics is shown schematically by the solid curves (*c*) for 8.1 GHz. After convolution with the intrinsic source diameter (*a*) an observer line-of-sight cutting across the structure at a distance *D* will produce a light curve (solid line in *d*) similar to that of 0954+658. In regions (i) and (iii) there will be three images; in (ii) one de-amplified image is present. At 2.7 GHz the refraction is stronger and the caustics (dotted lines in *c*) form closer to the plasma lens. The increase in θ_{int} (*a*) gives a light curve (*d*) with two broad maxima shown by the dotted line.

$\chi \approx f/C \approx 0.02$, so there are $\sim \chi \eta (R^2/\eta a^2) \approx 10^{11}$ sheet facets per remnant, containing ≤ 0.01 of the remnant mass.

An average line of sight through an old remnant will pass through $\sim \chi \eta^2 \approx 100$ sheets face-on. If these are uniform slabs, the associated scattering angle at $\sim$ metre wavelengths will be very small. More realistically, however, the sheets will have some transverse structure; for example, the density might have a corrugated modulation of width $\sim a$ (which could be organized by magnetic fields). With our fiducial values, the incidence of remnant-crossing lines of sight is then consistent with the oscillations seen $\sim 10\%$ of the time in pulsar dynamic scintillation spectra. These oscillations are interpreted as the beating between two or more sub-images of the pulsar^{3,10,11}; the sub-image splitting angles seen are a few μ arc s at 0.4 GHz, too large to be caused by the general scattering medium. For an average remnant-crossing line of sight, however, the scattering of a fraction $\sim 1/\eta$ of the corrugated face-on sheets adds incoherently in a given transverse dimension, giving a mean scattering angle $\sim (\eta \chi)^{1/2} (\theta_{ref}/\eta)$. With observed angles for our typical caustic event and fiducial parameters, this gives refraction angles of 1–10 μ arc s at $\leq$ GHz-observing frequencies, appropriate to the periodic frequency drifts and strong pulsar focusing mentioned above. We thus have a unified picture of thin sheet-like lensing structures causing the dramatic caustic events at $\sim$ cm wavelengths when seen edge-on, and the more common sub-GHz pulsar modulations when they intersect the line of sight at average inclinations.

As an alternative model for the production of the caustics, we can suppose that the lenses are confined magnetically, most

plausibly in twisted magnetic ropes or filaments of diameter $\sim 10^{14}$ cm. At this scale, a helical field of $\sim 5 \times 10^{-4}$ G allows confinement of densities $\sim 10^4$ cm $^{-3}$ where interstellar ultraviolet ionization and heating can balance radiative cooling¹². Here $\eta \sim 1$ and the required density of $\sim$ pc-length flux tubes is ~ 1 pc $^{-2}$. These ropes could be spun off by solar-type stars, reconnecting and combining as they cross one another. Filamentary structure, albeit on larger scales is observed in the galactic centre¹³.

These structures alone will not dominate the total energy, line radiation, pulsar dispersion measure and low-latitude intense scattering of the ISM. But they have far-reaching implications for very-long-baseline interferometry (VLBI), variability of compact radio sources, and precision timing of pulsars, although the modest duty cycle means that these events will be infrequently seen for a given source. The most extraordinary caustic events (such as the edge-on event seen from 0954+658) have associated splitting angles that are resolvable with intercontinental VLBI. Near the caustic, variation of the sub-images with time and observing frequency will be rapid. This variation of the bright, elongated sub-images can mimic a large superluminal motion in a 'core-jet' source, although such motion will be unrelated to the large-scale source structure and can be distinguished by its strong frequency dependence. Sustained imaging (with moderate bandwidth at several frequencies) during a caustic event, feasible with the Very-Long-Baseline Array should yield the detailed profile of column density through the refractor. More generally, further studies with VLBI and intensity monitoring will yield important information on the interstellar medium and this component of interstellar seeing, which can be an important limitation for Quasar. Because single perturbation, small- α events highlight the caustic variation, the most discerning observations will be at high (for example, ≥ 2 GHz) frequencies and moderate to high galactic latitudes.

A number of studies have addressed the influence of the distributed scattering component of the interstellar medium on pulsar arrival times.^{14–16} Arrival-time perturbations induced by strong refraction events include those due to dispersion measure (DM) fluctuations, geometrical pathlength variations and position errors in the timing fit. For a typical caustic event the DM and path length variations will dominate the arrival time perturbation, giving $\delta t \sim 0.4 \nu_{\text{GHz}}^2$ ms (or $\delta \text{DM} \sim 0.1$ pc cm $^{-3}$); if sheet structures are responsible face-on events may be detectable in millisecond pulsar (for example PSR1937+21) timing at $\delta t \sim 4 \mu$ s¹⁷. Rotation measure changes may also be detectable in polarization monitoring programmes^{18–20}. For standard parameters and a magnetic field $\sim 10^{-5}$ B $_{-5}$ G, position angle variations of ~ 0.1 B $_{-5}$ rad will be seen at 1 GHz in caustic events. If the field is important in organizing the structures, then even larger variations with characteristic signatures may be seen.

The observations of Fiedler *et al.* along with low-frequency variability of quasars and multiple imaging of pulsars point to high-density localized ISM electron density fluctuations in a dense ionized medium (DIM) on scales of $\sim 10^{14}$ cm which can individually dominate the refractive scattering due to the general ISM along a given line of sight. The proposed observations should provide extensive probes of this component of the ISM and clean tests of the models outlined above.

We thank Ralph Fiedler for introducing us to these observations and acknowledge valuable discussions with Don Backer, Marshall Cohen, Shri Kulkarni and George Rybicki. Support for this work was provided by the NSF and the National Astronomy and Ionosphere Center, operated by Cornell University under contract with the NSF. We also acknowledge fellowships from the Alfred P. Sloan Foundation (J.C.) and the Fannie and John Hertz Foundation (R.W.R.).

Received 13 March; accepted 7 May 1987.

1. Fiedler, R. L., Dennison, B., Johnston, K. J. & Hewish, A. *Nature* (submitted).
2. Romani, R. W., Narayan, R. & Blandford, R. D. *Mon. Not. R. astr. Soc.* **220**, 19–49 (1986).

3. Cordes, J. M. & Wolszczan, A. *Astrophys. J.* **307**, L27-L31 (1986).
4. Goodman, J. J., Romani, R. W., Blandford, R. D. & Narayan, R. *Mon. Not. R. astr. Soc.* (submitted).
5. Pearson, T. J. & Readhead, C. S. *Astrophys. J.* **234**, 61-81 (1981).
6. McKee, C. F. & Ostriker, J. P. *Astrophys. J.* **218**, 148-169 (1977).
7. Hamilton, P. A., McCulloch, P. M., Manchester, R. N., Ables, J. G. & Komesaroff, M. M. *Nature* **265**, 224-225 (1977).
8. Hester, J. J. *Astrophys. J.* **314**, 187-202 (1987).
9. Hester, J. J., Parker, R. A. R. & Dufour, R. J. *Astrophys. J.* **273**, 219-242 (1983).
10. Roberts, J. & Ables, J. *Mon. Not. R. astr. Soc.* **201**, 1119-1138 (1982).
11. Hewish, A., Wolszczan, A. & Graham, D. *Mon. Not. R. astr. Soc.* **213**, 167-179 (1985).
12. Reynolds, R. J. *Astrophys. J.* **282**, 191-196 (1984).
13. Yusef-Zadeh, F., Morris, M. & Chance, D. *Nature* **310**, 557-561 (1984).
14. Armstrong, J. W. *Nature* **307**, 527-528 (1984).
15. Blandford, R. D., Narayan, R. & Romani, R. W. *J. Astr. Astrophys.* **5**, 369-388 (1984).
16. Cordes, J., Pidwerbetsky, A. & Lovelace, R. V. *Astrophys. J.* **310**, 737-767 (1986).
17. Rawley, L. thesis, Princeton Univ. (1986).
18. Aller, H. D., Aller, M. & Hodge, P. E. *Astr. J.* **86**, 325-339 (1981).
19. O'Dea, C. P., Dent, W. A., Balonek, T. J. & Kapitzky, J. E. *Astr. J.* **88**, 1616-1625 (1983).
20. Caswell, J. L. & Lerche, I. *Mon. Not. R. astr. Soc.* **187**, 201-216 (1979).

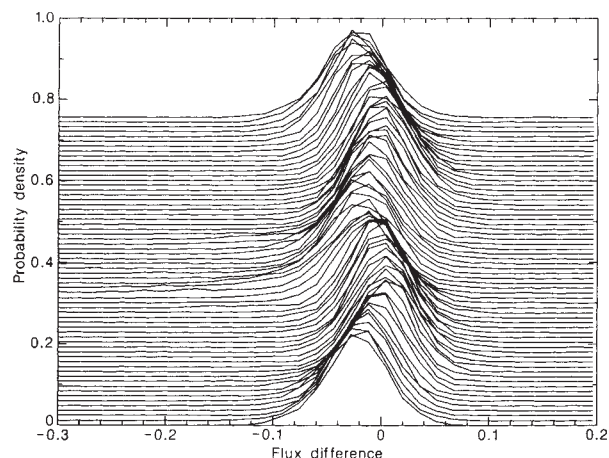


Fig. 1 Red-green flux difference distributions for 1985 solar-limb observations. Middle curves are for the equatorial bins and top and bottom traces are from polar observations. Distributions from successive latitude bins are displaced vertically.

Evidence of global circulation currents from solar-limb temperature variations

J. R. Kuhn*, K. G. Libbrecht† & R. H. Dicke‡

* Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824-1116, USA

† Solar Astronomy, California Institute of Technology 264-33, Pasadena, California 91125, USA

‡ Department of Physics, Princeton University, Princeton, New Jersey 08544, USA

The temperature distribution in a turbulent rotating photosphere is non-spherical. Dimensional arguments for the Sun suggest that such a temperature modulation may have an amplitude $\Delta T \approx T v^2 / \phi \approx 0.1$ K, where T is an average temperature (5,700 K) and v^2 / ϕ is the ratio of the rotational kinetic and potential energy density of the photosphere. Detailed calculations¹⁻³ generally support this expectation. Here we report new observations that should help to understand the solar global dynamics problem. From about 1,400 h of solar-limb data obtained during the summers of 1983-85 we find that the solar-limb temperature variation is not spherically symmetric and is ~ 1 K. Our results also indicate that the limb temperature departs from its expected $l = 2$ spatial harmonic form and has, at most, a weak dependence on solar cycle.

If the latitudinal differential rotation is 'driven' by angular momentum transport throughout the convection zone, then meridional flows and large-scale convection ('giant cells') should be present. A corresponding modulation in the convective heat flux due to these giant cells is then also expected. Analytical solutions² for the driving Reynolds stress associated with large convective cells must assume a spatial scale for the convective motion from which 'mean field' calculations generally yield a lowest-order temperature or velocity field. On the other hand, the best numerical simulations³ lack the resolution to model the solar envelope from the base of the convection zone out to the photosphere. Observations of the spatial and amplitude scale of the mean photospheric velocity and temperature fields are needed. A time base of several years is also required to understand the coupling with the solar dynamo and activity cycle.

The analysis of sunspot motions and Doppler shift data by Ribes *et al.*⁴ and Snodgrass and Howard⁵ now seem to suggest that the mean velocity field has two cycles per hemisphere. Although giant cells have not been unambiguously observed, these and earlier temperature data⁶ may be interpreted to suggest that downdraught regions exist near mid-latitudes between two (perhaps axisymmetric) cells. The regions of larger or smaller rotation rate that Snodgrass and Howard observe could be understood as the result of the Coriolis acceleration of the meridional circulation current from such cells.

Direct spectroscopic meridional-velocity measurements are very difficult to interpret, but the sunspot drift analysis of Ribes

*et al.*⁴ seems to show evidence of meridional motions of the above form with transverse velocities near 30 m s^{-1} . In contrast, the azimuthal flow velocities seen by Snodgrass and Howard were only a few m s^{-1} . One might have expected similar azimuthal velocities, especially if the cells are elongated in the meridional direction by rotation². Such conclusions are delicate because sunspot drift velocities are not directly comparable with the photospheric velocity observations; nevertheless there does not appear to be an azimuthal velocity field structure indicative of a pattern of convective cells. The flow pattern seems more strongly associated with updraught and downdraught latitude zones. Ribes *et al.* called this a convective 'roll' structure in their sunspot drift analysis.

The exponential vertical density variation could give rise to much larger transverse than radial velocities for large-scale surface flows. For example, a quasi-stationary flow with the approximate velocity amplitude and latitude structure observed by Ribes *et al.* might have vertical velocity amplitudes near the photosphere of only 1 cm s^{-1} . Although such velocities will be very difficult to observe directly, the corresponding local convective flux perturbation could be measurable. Motivated by the mixing length theory we might expect a relative convective flux perturbation $\Delta F / F \approx \Delta u / u$, where $\Delta u / u$ is the fractional perturbation of vertical eddy velocity. In this case $\Delta u / u \approx 10^{-4}$. As we note below the corresponding temperature perturbations at the level of $\Delta T / T \approx 10^{-3}$ or 10^{-4} are observable. The best probe of such a flow structure may be temperature measurements like those described here.

We previously reported⁶ evidence for a latitude-dependent limb-temperature variation from data obtained in 1983. These and later data, obtained in 1984 and 1985 (ref. 7), were obtained with the Princeton solar distortion telescope. This specialized instrument simultaneously samples light from the solar limb in two colour bands. The integrated flux is taken at three different limb exposures ranging from ~ 5 to 20 arc s from the limb. The telescope combines the light flux from two regions separated by 180° with a resolution of 256 angular bins around the limb. A statistical technique for analysing the colour difference signal at each bin and radial limb exposure was developed, and yielded: (1) the colour probability distribution function for the background photosphere, (2) a corresponding facular distribution, (3) the latitude distribution of the facular colour contributions, and (4) an effective colour shift function that describes the latitude-dependence of the background photospheric colour distributions. This last function could be directly related to an effective temperature variation of the photosphere and provided direct evidence of a latitude-dependent limb-temperature variation. Faculae were often detected near the limb because the

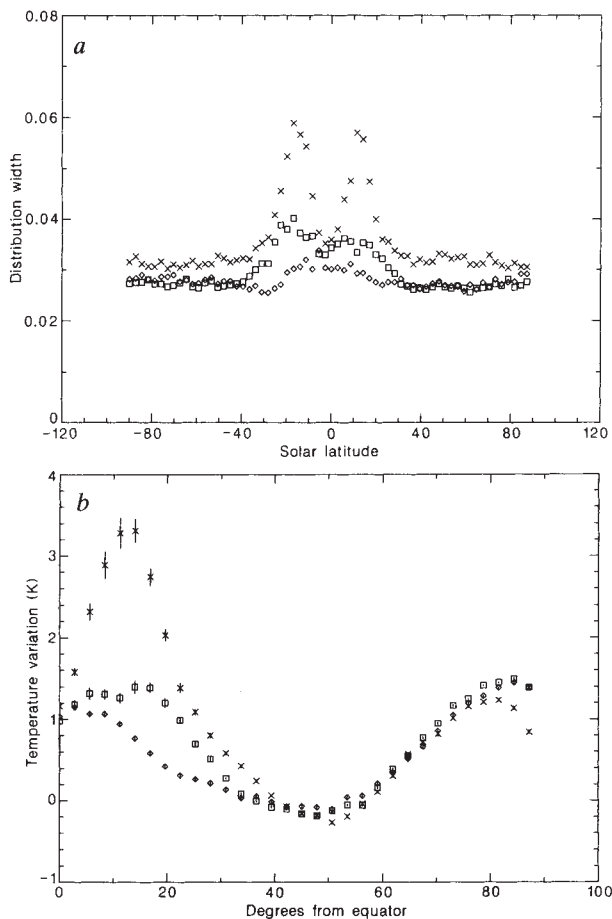


Fig. 2 *a*, Gaussian half-width of Fig. 1 distributions for 1983–85 data. Crosses, 1983; squares, 1984; diamonds, 1985. *b*, Latitudinal limb temperature variation corresponding to the distribution colour shifts for 1983–85 data. Symbols as in *a*.

Sun was closer to its activity maximum in 1983. This relatively complex analysis was required to distinguish the effects of unresolved faculae from the variation in limb temperature that was observed.

Data obtained during 1984 and 1985 have a much smaller facular component, too small for the analysis described above to resolve the photospheric and facular distribution functions. Thus, having established the relatively minor importance of the faculae to the 1983 temperature measurements, we adopted a simpler analysis of the colour distribution data reported here. The technique has the advantage that it can be uniformly applied to the 1983–85 data sets.

Figure 1 shows the distribution functions of the difference in solar limb flux between the 'red' (800 nm) and 'green' (525 nm) observations obtained with the largest limb exposure (19.6 arc s in the red band) in 1985. As described in ref. 6 these measurements are expressed in terms of the equivalent solar limb displacement (in arc s) that would produce the observed brightness signal with a given exposed limb. The brightness is measured with respect to the circular mean, so that when a difference is taken, using two simultaneous colour measurements, the results are insensitive to the geometrical limb shape and most photometric errors associated with the telescope and sky⁶. Successive vertically displaced curves in Fig. 1 are the distributions from different latitude bins. The middle traces are from the equatorial region and the upper and lower curves describe the solar polar regions. Two notable features of Fig. 1 are: (1) the distributions show very little variation in width except near the low-latitude active regions, (2) there is a systematic colour shift of the curves from the pole to equator. This is quantified in Fig. 2, which shows the gaussian width of the largest limb-exposure data from

1983–85 and the local temperature change required to produce the observed shift in the colour distributions. Error bars (1σ) are derived from the fits to the distributions.

The results in Fig. 2*a* suggest that the temperature dip near $\pm 50^\circ$ solar latitude is not due to an otherwise uniformly bright photosphere with cool spots in this latitude band rotating across the limb. If this were so, the width of the colour distribution would be broader near mid-latitudes. As it is, the shift in the distributions with latitude is larger than their half-widths, which are constant outside the active latitude bands with an uncertainty of $\sim 10\%$. Thus, whatever is responsible for the limb temperature variation has an essentially global spatial scale, with a greater latitudinal than azimuthal temperature structure.

Figure 2*b* shows only a weak 3-yr time dependence in the measured temperatures. The large variations at latitudes within about 35° of the equator are due, at least in part, to the residual effects of faculae on the shift of the distributions. Yet it is clear from these data that any latitude shift in the temperature minima near 50° (between 1983 and 1985) is small.

Data obtained with smaller limb exposures from 1984 and 1985 show a similar temperature structure but with an effective temperature amplitude $\leq 50\%$ larger at latitudes greater than about 35° . The latitude of the temperature minima also varies by $\leq 10^\circ$ between the small limb-exposure data sets. We believe that the trend towards decreasing temperature perturbation with increasing limb exposure at high latitudes is partly a result of our simple model of the effects of a temperature change on the extreme limb-darkening profile. We have assumed that the flux perturbation is limb-distance independent and has a black-body spectral profile. Better model solar-atmosphere calculations⁸ suggest that this may be only approximately correct. Such calculations also indicate that a significant decreasing variation in effective temperature with decreasing limb exposure would be seen if the flux variation is caused by latitudinal variations in the local convective mixing length or small-scale granule energy transport parameters. At high latitudes, where facular contamination is negligible, this trend can be ruled out, suggesting that these limb observations are not related to a simple latitudinal variation in granule properties. It is notable that local properties of granules can be measured without spatially resolving them from high spectral resolution observations of asymmetry in Fraunhofer line profiles. No such latitude dependence has been detected by spectroscopic techniques either^{9,10}.

Given a convection-zone depth of $d \approx 0.3R_\odot$ it is not surprising that the transverse scale of the largest convective cell structure should be $\sim 2d/R_\odot \approx 40^\circ$. In fact Glatzmaier³ does see this '2 cell per hemisphere' behaviour in his numerical simulations, but he also finds that the kinetic energy of the non-axisymmetric motions is much larger than the meridional flows. This appears to contradict both the temperature and velocity data, which show a weaker azimuthal structure. Perhaps a global toroidal magnetic field near the bottom of the convection zone, as discussed by Parker¹¹, can also account for the largely axisymmetric temperature modulations that we have observed.

This research was supported in part by the NSF and a Sloan foundation grant to J.R.K. We thank R. Stein and P. Wilson for carefully reading the manuscript.

Received 24 March; accepted 8 May 1987.

1. Dicke, R. H. & Goldenberg, H. M. *Nature* **214**, 1294–1296 (1967).
2. Durney, B. R. & Spruit, H. C. *Astrophys. J.* **234**, 1067–1078 (1979).
3. Glatzmaier, G. *Astrophys. J.* **291**, 300–307 (1985).
4. Ribes, E., Mein, P. & Mangeney, A. *Nature* **318**, 170–171 (1985).
5. Snodgrass, H. B. & Howard, R. *Sol. Phys.* **95**, 221–228 (1985).
6. Kuhn, J. R., Libbrecht, K. G. & Dicke, R. H. *Astrophys. J.* **290**, 758–766 (1985); erratum, *Astrophys. J.* (in the press).
7. Kuhn, J. R. in *The Internal Solar Angular Velocity: Theory and Observations* (eds Durney, B. & Sofia, S.) (Reidel, Dordrecht, in the press).
8. Petro, L. D., Foukal, P. V. & Kurucz, R. L. *Sol. Phys.* **98**, 23–38 (1984).
9. Kuhn, J. R. in *Small Scale Dynamical Processes in Quiet Stellar Atmospheres* (ed. Keil, S.) 381–388 (National Solar Observatory, Sunspot, New Mexico, 1984).
10. Cavallini, F., Ceppatelli, B. & Righini, A. *Astr. Astrophys.* **150**, 256–264 (1985).
11. Parker, E. in *The Internal Solar Angular Velocity: Theory and Observations* (eds Durney, B. & Sofia, S.) (Reidel, Dordrecht, in the press).

The structure of the non-superconducting phase $\text{La}_3\text{Ba}_3\text{Cu}_6\text{O}_{14+x}$ and its relation to the high- T_c superconductor $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$

W. I. F. David*, W. T. A. Harrison*, R. M. Ibberson*, M. T. Weller†, J. R. Grasmeyer† & P. Lanchester‡

* Neutron Division, Rutherford Appleton Laboratory, Chilton, Didcot OX11 0QX, UK

Departments of Chemistry† and Physics‡, University of Southampton, Southampton SO9 5NH, UK

The compound $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ has recently received intense interest because of its high-temperature ($T_c \approx 95$ K) superconducting properties¹. Shortly after its discovery, many authors independently confirmed that the structure could be regarded as an orthorhombically distorted tripled perovskite²⁻⁵. We have investigated $\text{La}_3\text{Ba}_3\text{Cu}_6\text{O}_{14+x}$ because it is chemically similar to $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ and because initial work⁶ indicated a poor structure determination based on a cell not dissimilar from $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$. Here we report time-of-flight neutron powder diffraction results which confirm that the structure of $\text{La}_3\text{Ba}_3\text{Cu}_6\text{O}_{14+x}$ is isomorphous with the tetragonal variant of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$; in particular, the copper coordination and calculated valencies of both compounds agree closely. The apparent contradiction of stoichiometries between these two phases is resolved by ordering of the large cations consistent with a formulation $\text{La}(\text{La}_{0.25}\text{Ba}_{0.75})_2\text{Cu}_3\text{O}_{7+\frac{1}{2}x}$. Stoichiometries for both quenched and oxygen-annealed materials are in excellent agreement with those obtained by Er-Rakho *et al.*⁶, yielding more than seven oxygens per formula unit. The excess oxygen is located in the planes containing the chains of CuO_4 planar groups in $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$. The present results extend the upper limit of oxygen stoichiometry from $\text{RA}_2\text{Cu}_3\text{O}_7$ to $\text{RA}_2\text{Cu}_3\text{O}_{7.2}$, but the materials studied here remain semiconducting down to 7 K (Fig. 1). Thus, although there is a remarkable structural stability from $\text{RA}_2\text{Cu}_3\text{O}_6$ (ref. 7) to $\text{RA}_2\text{Cu}_3\text{O}_{7.2}$, superconductivity appears to be confined to the range $\text{RA}_2\text{Cu}_3\text{O}_{6.5}$ to $\text{RA}_2\text{Cu}_3\text{O}_7$.

Neutron diffraction data were collected on the ISIS high-resolution powder diffractometer (HRPD⁸) for two samples of $\text{La}_3\text{Ba}_3\text{Cu}_6\text{O}_{14+x}$ prepared following ref. 6. The first sample ($x \approx 0.43$) was sintered in air at 950 °C and annealed at 350 °C in oxygen: the second ($x \approx 0.11$) was sintered at 950 °C and quenched to room temperature. All diffraction lines could be indexed on the basis of a tripled perovskite cell; no evidence was found for the superlattice cell doubling proposed by Er-Rakho *et al.*⁶ and no orthorhombic distortion was observable. Refinements were thus performed in the tetragonal space group $P4/mmm$ based on an $a \times a \times 3a$ perovskite structure. Oxygen atoms were positioned at the sites allowed in a fully oxidized perovskite (18 oxygens per formula unit) and their occupancies were refined. The oxygen vacancy ordering corresponding to the tetragonal variant of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ was rapidly obtained. Refined oxygen stoichiometries (Table 1) are in excellent agreement with ref. 6. The oxygen in excess of 14 formula units is located interstitially in the layer containing the CuO_4 square planar groups (see ref. 3), indicating the presence of CuO_5 square pyramids and/or CuO_6 octahedra. Refinement of the lanthanum and barium occupancies in the two crystallographically distinct large-cation sites indicates full occupation by La of the site at $(1/2, 1/2, 1/2)$ and a 75% Ba:25% La occupation of the site at $(1/2, 1/2, 0.18)$, and is confirmed by bond valency calculations. Thus, by analogy with $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$, the structure may be formulated as $\text{La}(\text{La}_{0.25}\text{Ba}_{0.75})_2\text{Cu}_3\text{O}_{7+\frac{1}{2}x}$: excess La^{3+} essentially dopes the Ba^{2+} site. Full details of the refinements, including selected bond distances, are listed in Tables 1 and 2. The final powder profiles for $\text{La}_3\text{Ba}_3\text{Cu}_6\text{O}_{14.39}$ are shown in Fig. 2.

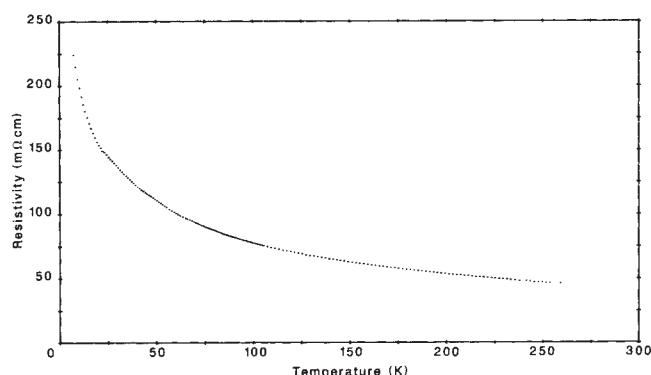


Fig. 1 The resistivity of annealed $\text{La}_3\text{Ba}_3\text{Cu}_6\text{O}_{14.39}$ as a function of temperature, showing typical semiconducting behaviour.

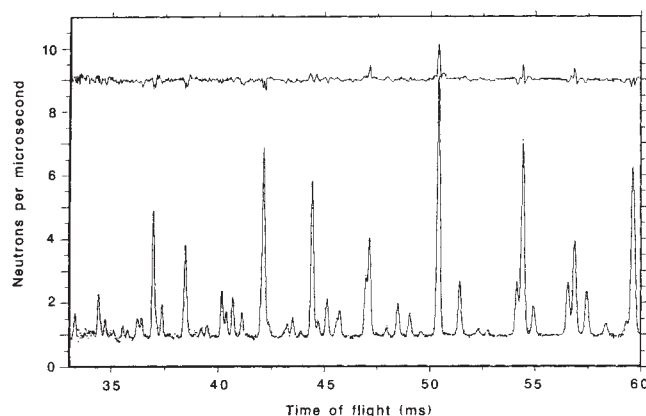


Fig. 2 A section of the observed and calculated diffraction patterns for annealed $\text{La}_3\text{Ba}_3\text{Cu}_6\text{O}_{14.39}$, showing the excellent profile fit obtained ($0.68 \text{ \AA} < d < 1.24 \text{ \AA}$ for this section of the pattern).

Assuming formal valences, the oxygen stoichiometries in $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ corresponding to our annealed and unannealed samples, respectively, are $7 - \delta = 6.945(30)$ and $6.810(50)$ —assuming identical overall copper charge. As the yttrium compounds with such stoichiometries are superconducting, it may be concluded that the average copper valence is not the sole criterion in determining superconductivity. Indeed, analysis of bond strengths (Table 2) shows that the division of valence between the two copper sites is very similar between the $\text{La}_3\text{Ba}_3\text{Cu}_6\text{O}_{14+x}$ phases and $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$. There are two distinct differences between $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ and $\text{La}(\text{La}_{0.25}\text{Ba}_{0.75})_2\text{Cu}_3\text{O}_{7+\frac{1}{2}x}$. (1) Lattice symmetry: orthorhombic versus tetragonal. The tendency towards a tetragonal lattice in the large-rare-earth compounds may be understood as follows. A large rare-earth ion at $(1/2, 1/2, 0.18)$ may be regarded as exerting a large internal pressure that is relieved most easily by elongating the $\text{Cu}(2)$ – $\text{Cu}(2)$ c -axis separation, which also shortens the $\text{Cu}(2)$ – $\text{O}(3)$ bond length (Table 2), compared with the corresponding bond in $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$. This expansion is limited and the remaining ‘pressure’ is most easily accommodated by increasing the separation (along the a -axis) between the CuO_4 square-planar chains rather than stretching the chains themselves (the b -axis). At a certain rare-earth size, the chain separation becomes equivalent to the lattice repeat of the chain. Little energy is then required to disorder the oxygens over the a - and b -axes: the compound thus becomes tetragonal. Recent results⁹ for $\text{LaBa}_2\text{Cu}_3\text{O}_7$ show that the CuO_4 square-planar chains may tetragonally disorder and yet still permit superconductivity. Thus tetragonality is not the sole reason for the lack of superconductivity in the $\text{La}_3\text{Ba}_3\text{Cu}_6\text{O}_{14+x}$ phases. (2) Doping of the barium site by lanthanum in $\text{La}(\text{La}_{0.25}\text{Ba}_{0.75})_2\text{Cu}_3\text{O}_{7+\frac{1}{2}x}$, along with excess oxygen located interstitially alongside the chains of CuO_4 planar groups to create isolated CuO_5 square pyramids and/or

Table 1 Final crystallographic data for the $\text{La}_3\text{Ba}_3\text{Cu}_6\text{O}_{14+x}$ phases

La₃Ba₃Cu₆O_{14.39} (annealed)						La₃Ba₃Cu₆O_{14.11} (unannealed)					
Tetragonal: space group <i>P4/mmm</i> (no. 123)						Tetragonal: space group <i>P4/mmm</i> (no. 123)					
<i>a</i> = 3.90018(8) Å		<i>b</i> = 3.90018(8) Å		<i>c</i> = 11.69265(25) Å		<i>a</i> = 3.90510(12) Å		<i>b</i> = 3.90510(12) Å		<i>c</i> = 11.71400(40) Å	
Atom	Wyckoff symbol	<i>x/a</i>	<i>y/b</i>	<i>z/c</i>	Site occupancy	Atom	Wyckoff symbol	<i>x/a</i>	<i>y/b</i>	<i>z/c</i>	Site occupancy
La	1d	1/2	1/2	1/2		La	1d	1/2	1/2	1/2	
(La _{0.25} Ba _{0.75})	2h	1/2	1/2	0.17928 (20)		(La _{0.25} Ba _{0.75})	2h	1/2	1/2	0.18266 (30)	
Cu(1)	1a	0	0	0		Cu(1)	1a	0	0	0	
Cu(2)	2g	0	0	0.34472 (13)		Cu(2)	2g	0	0	0.34644 (23)	
O(1)	2f	0	1/2	0	0.598 (16)	O(1)	2f	0	1/2	0	0.529 (24)
O(2)	4i	0	1/2	0.36402 (15)		O(2)	4i	0	1/2	0.36503 (25)	
O(3)	2g	0	0	0.15816 (22)		O(3)	2g	0	0	0.15730 (40)	
Temperature factors						Temperature factors					
Atom	<i>B</i> ₁₁ (Å ²)	<i>B</i> ₂₂ (Å ²)	<i>B</i> ₃₃ (Å ²)			Atom	<i>B</i> ₁₁ (Å ²)	<i>B</i> ₂₂ (Å ²)	<i>B</i> ₃₃ (Å ²)		
La	0.44 (5)	0.44 (5)	1.78 (13)			La	0.18 (9)	0.18 (9)	2.08 (22)		
(La _{0.25} Ba _{0.75})	1.35 (7)	1.35 (7)	1.95 (14)			(La _{0.25} Ba _{0.75})	1.48 (13)	1.48 (13)	1.68 (21)		
Cu(1)	2.27 (9)	2.27 (9)	0.67 (13)			Cu(1)	2.50 (15)	2.50 (15)	0.38 (22)		
Cu(2)	0.57 (4)	0.57 (4)	0.61 (9)			Cu(2)	0.54 (6)	0.54 (6)	0.47 (14)		
O(1)	11.4 (5)	3.4 (3)	0.93 (27)			O(1)	11.4 (1.1)	3.4 (6)	2.1 (5)		
O(2)	1.10 (6)	0.46 (5)	1.83 (5)			O(2)	0.76 (10)	0.36 (9)	2.14 (9)		
O(3)	4.15 (11)	4.15 (11)	0.54 (13)			O(3)	4.7 (2)	4.7 (2)	0.65 (24)		
<i>R</i> _E = 5.0%; <i>R</i> _{WP} = 7.6%; <i>R</i> _P = 6.8%; <i>R</i> _N = 7.1%; <i>N</i> − <i>P</i> + <i>C</i> = 6,211; χ^2 = 2.3						For a definition of the <i>R</i> -factors see ref. 3.					
Calculated stoichiometry = La ₃ Ba ₃ Cu ₆ O _{14.39(6)}						<i>R</i> _E = 6.4%; <i>R</i> _{WP} = 8.4%; <i>R</i> _P = 7.6%; <i>R</i> _N = 8.9%; <i>N</i> − <i>P</i> + <i>C</i> = 6,211; χ^2 = 1.70					
Average particle size = 1,010 Å (determined from refined peak width parameters)						Calculated stoichiometry = La ₃ Ba ₃ Cu ₆ O _{14.11(10)}					
						Average particle size = 650 Å					

Table 2 Bond lengths ($\text{\AA}$) and copper valences for the $\text{La}_3\text{Ba}_3\text{Cu}_6\text{O}_{14+x}$ phases

Bond	Annealed ($x = 0.39$)		Unannealed ($x = 0.11$)		$\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$	
	Length	Number	Length	Number	Length	Number
Cu(1)-O(1)	1.9501	$\times 2.39$	1.9526	$\times 2.11$	1.942	$\times 2$
Cu(1)-O(3)	1.8493	$\times 2$	1.8424	$\times 2$	1.843	$\times 2$
Cu(2)-O(2)	1.9631	$\times 4$	1.9647	$\times 4$	1.929/1.958	$\times 2$
Cu(2)-O(3)	2.1814	$\times 1$	2.2161	$\times 1$	2.306	$\times 1$
La-O(2)	2.5160	$\times 8$	2.5124	$\times 8$	2.387/2.402	$\times 4$
La/Ba-O(1)	2.8631	$\times 2.39$	2.8967	$\times 2.11$	2.827	$\times 4$
La/Ba-O(2)	2.9102	$\times 4$	2.8943	$\times 4$	2.986/2.949	$\times 2$
La/Ba-O(4)	2.7689	$\times 4$	2.7773	$\times 4$	2.741	$\times 2$

The bond lengths in the third column are those for the corresponding bonds in $\text{YBa}_2\text{Cu}_3\text{O}_{6.85}$ (ref. 3). The La site corresponds to Y, the La/Ba site to Ba. The partial bond orders are related to the site occupancy of O(1) for the $\text{La}_3\text{Ba}_3\text{Cu}_6\text{O}_{14+x}$ phases, for example, in $\text{La}_3\text{Ba}_3\text{Cu}_6\text{O}_{14.39}$ Cu(1) is coordinated on average by 2.39 O(1) throughout the structure.

Assuming:	Calculated valences (after ref. 3)		
	Cu^{2+}	Cu^{3+}	Σ_s
annealed Cu(1) site	2.40	2.70	2.57
annealed Cu(2) site	2.04	2.18	2.04
non-annealed Cu(1) site	2.29	2.58	2.40
non-annealed Cu(2) site	2.01	2.14	2.01
$\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ Cu(1) site	2.27	2.56	2.38
$\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ Cu(2) site	2.01	2.16	2.01

Σ_2 and Σ_3 are defined to be respectively the copper valences assuming Cu^{2+} and Cu^{3+} bond strength parameters, then Σ_s , given by the equation $(\Sigma_s - \Sigma_2)/(\Sigma_3 - \Sigma_2) = \Sigma_2 - 2$ represents an internally consistent measure of the valence for each copper site. Thus, for all three phases, the bond valence calculations indicate that Cu(1) has mixed $\text{Cu}^{2+}/\text{Cu}^{3+}$ character whilst Cu(2) is predominantly Cu^{2+} in character.

CuO_6 octahedra. These oxygen localizations, which may only occur at oxygen stoichiometries of >7 , may lead to the localization and trapping of electrons in the copper $d_{x^2-y^2}$ band associated with the chains of CuO_4 square-planar groups. This static perturbation may then inhibit the movement of electrons along the chains, suppressing superconductivity and indicating the importance of non-disrupted strings of CuO_4 square-planar groups in superconducting $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$. This would suggest that superconductivity in these phases is possible only when the overall oxygen stoichiometry is <7 .

Received 29 June; accepted 1 July 1987.

- Wu, M. K. *et al. Phys. Rev. Lett.* **58**, 908-911 (1987).
- Beno, M. A. *et al. Appl. Phys. Lett.* (in the press).
- David, W. I. F. *et al. Nature* **327**, 310-312 (1987).
- Capponi, J. J. *et al. Europhys. Lett.* **3**, 301-308 (1987).
- Siegrist, T., Sunshine, S., Murphy, D. W., Cava, R. J. & Zahurak, S. M. *Phys. Rev. B* **35**, 7137-7140 (1987).
- Er-Rakho, L., Michel, C., Provost, J. & Raveau, B. *Solid St. Chem.* **37**, 151-156 (1981).
- Bordet, P., Chailout, C., Capponi, J. J., Chenavas, J. & Marezio, M. *Nature* **327**, 687-689 (1987).
- Johnson, M. W. & David, W. I. F. *Rutherford Appleton Lab. Rep.* No 85/112 (1985).
- Michel, C. *et al. C. r. hebdom. Séanc. Acad. Sci. Paris* **304**, 1169-1172 (1987).

Geomagnetic intensity as evaluated from ancient Chinese pottery

Wei Quing-Yun*, Zhang Wei-Xi*, Li Dong-Jie*,
M. J. Aitken†, G. D. Bussell† & M. Winter†

* Institute of Geophysics, Academy of Sciences, Beijing, China

† Research Laboratory for Archaeology, Oxford University,
6 Keble Road, Oxford OX1 3QJ, UK

Past values for the geomagnetic intensity may be obtained by laboratory analysis of the thermoremanent magnetization carried by clay baked in ancient times. From global averages¹⁻³ of such determinations it is commonly accepted that the intensity in any given region went through a broad maximum about 2,000 years ago, reaching a level ~50% higher than at present. Here we present results obtained from a wide range of Chinese pottery, spanning the interval from 4000 BC to the present, indicating that the field behaviour was more complex. The intensity was high between 1500 and 1000 BC and again in the first half of the first millennium AD. Comparison with results reported^{4,5} for Western Asia, Egypt and Crete suggests that these high values are due to non-dipole disturbances in the geomagnetic field, consistent with long-term records⁶⁻⁸ of the cosmogenic radioisotopes ¹⁴C and ¹⁰Be.

Samples were obtained from museums and excavations in various provinces as indicated in Fig. 1. Although about 450 were collected only 70 had sufficiently satisfactory magnetic characteristics for results of significance to be obtained; these are given in Table 1 along with the date spans used. From 2200 BC onwards the commonly accepted dynastic dates have been used in most cases; from 840 BC these are essentially based on continuous archives anchored in time by recorded astronomical observations. Earlier than 2200 BC radiocarbon dating is used; there are only 20 samples in this category.

Sampling was by means of a diamond-impregnated corer, 3 mm × 3 mm cylinders of baked clay thereby being obtained. Thellier-type⁹ determinations were made using a superconducting quantum interference device (SQUID) cryogenic magnetometer as previously described elsewhere^{10,11}; as usual, reliability was assessed on the basis of magnetic characteristics, several cores being measured from each sample. In addition to pottery, some bricks and tiles were used. From the 70 porcelain or protopcelain samples that were tried, useful results were obtained from only three; most of these samples were too weak (with less than a total moment of 2.5×10^{-9} Am²) for satisfactory measurement with the 3-mm core size. Figure 2 shows the results for samples having good reliability (grades 1 and 2), samples of poor reliability (grade 3) being omitted. The vertical error limits correspond to ±6%, which is our estimate of the average standard deviation from the true value of the result for a sample of good reliability; this estimate was obtained from comparisons with observatory-based values using samples fired in the last 150 years. The results are expressed as the ratio between the ancient field F^A and the field F^D at the site due to an axial geocentric dipole of present day strength (8×10^{22} Am²). This is a normalized version of the virtual axial dipole moment (VADM) presentation suggested by Barton *et al.*¹², and, similarly it achieves first-order removal of latitude dependence (that due to an axial-centred dipole). As can be seen from Fig. 1, except in two cases, the results of Fig. 2 refer to samples fired on sites between latitudes 30° and 40° N and longitudes 105° and 120° E, having a centroid around 33° N, 113° E.

Although the high values for the intensity ratio around 1000 BC parallel those found for Western Asia⁴, Egypt⁴ and Crete⁵ at that time, the fall to values of less than unity for several centuries following 500 BC does not, any minimum in the values for the latter regions occurring after the beginning of the Christian era. Thus one interpretation of the data sets is in terms of



Fig. 1 Map of China showing localities from which samples were obtained (shaded). Localities for the samples included in Fig. 2 are indicated by letters: H (Henan), J (Jiangsu), S (Shanxi), Gd (Guangdong) and Gu (Guangxi).

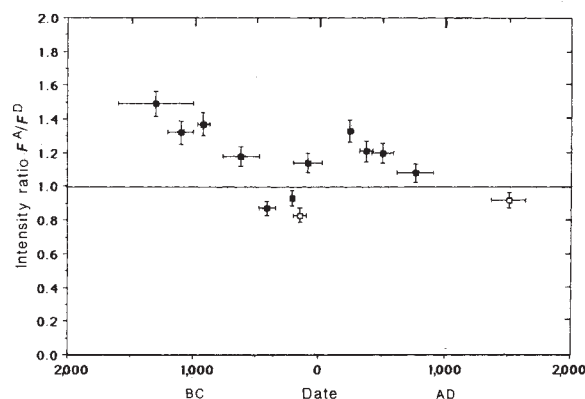


Fig. 2 Intensity ratio, F^A/F^D , for samples having good reliability, where F^A is the ancient field intensity and F^D is the field intensity at the sample locality due to an axial geocentric dipole of present-day strength (8×10^{22} Am²). These samples came from provinces marked S, H and J in Fig. 1, except for the samples at 200–100 BC and AD 1370–1640, which came from provinces Gd and Gu respectively.

a non-dipole disturbance that drifted westward through approximately 110° of longitude in about half a millennium. It is interesting to recall that the rate of westward drift of the non-dipole field derived by Bullard *et al.*¹³ from observatory data for the first half of the twentieth century was 0.2° per year. The interpretation in terms of a non-dipole disturbance is consistent with the long-term records⁶⁻⁸ of the cosmogenic radioisotopes ¹⁴C and ¹⁰Be which, in the time period concerned, show no features indicative of the reduced atmospheric concentrations which would occur following the decrease in stratospheric cosmic ray flux that would ensue from an enhancement of the dipole component of the geomagnetic field—but not from an enhancement of the non-dipole component.

This work was undertaken as a joint project of the two institutions concerned, sponsored by the Chinese Academy of Science and the Royal Society of London; there was also support from the NERC. We thank the archaeologists and museum curators who have provided samples. We thank J. Shaw for confirmatory measurements on some of the samples, P. Alcock

Table 1 Magnetic characteristics and date spans of the samples

Sample	Site	Latitude (deg.)	Longitude (deg.)	Dynasty	Date	F^A (μ T)	Intensity ratio	Grade
1. Tile	Beijing region	40.0	116.2	Qing	AD1796-1820	54.8	1.07	3
2. Tile	Beijing region	40.0	116.2	Qing	AD1796-1820	49.8	0.97	3
3. Porcelain	Xinghua Kiln, Guilin, Guangxi prov.	25.3	110.2	Ming	AD1368-1644	44.8	1.05	3
4. Porcelain	Xinghua kiln, Guilin, Guangxi prov.	25.3	110.2	Ming	AD1368-1644	39.5	0.92	2
5. Porcelain	Xinghua kiln, Guilin, Guangxi prov.	25.3	110.2	Ming	AD1368-1644	43.3	1.01	3
6. Sherd	Guilin region, Guangxi prov.	25.3	110.2	Ming	AD1368-1450	44.8	1.05	3
7. Sherd	Xian region, Shanxi prov.	34.4	109.2	Yuan	AD1271-1368	46.7	0.97	3
8. Brick	Luoyang region, Henan prov.	34.7	112.7	Northern Song	AD1061-1160	42.3	0.91	3
9. Sherd	Luoyang, Henan prov.	34.7	112.7	Tang	AD618-907	51.9	1.08	2
10. Porcelain	Mengyang, Zhengzhou, Henan prov.	34.7	113.6	Tang	AD738-798	36.1	0.80	3
11. Sherd	Xian region, Shanxi prov.	34.4	109.2	Tang	AD738-798	51.3	1.11	3
12. Brick	Chaoyang region, Jilin prov.	42.7	126.1	Early Tang	AD618-738	56.6	1.05	3
13. Brick	Henan prov.			Sui Tang	AD581-618	68.4	1.42	3
14. Sherd	Xian region, Shanxi prov.	34.4	109.2	South/North	AD420-589	58.1	1.20	2
15. Brick	Luoyang region, Henan prov.	34.7	112.7	Northern Wei	AD471-520	67.1	1.42	3
16. Sherd	Luoyang region, Henan prov.	34.7	112.7	Northern Wei	AD361-534	69.0	1.43	3
17. Sherd	Luoyang region, Henan prov.	34.7	112.7	Jin	AD317-420	58.1	1.21	1
18. Brick	Luoyang region, Henan prov.	34.7	112.7	Western Jin	AD265-316	64.4	1.38	3
19. Sherd	Beijing region	40.0	116.2	Three Kingdoms	AD220-265	59.9	1.33	2
20. Sherd	Hanjing region, Jiangsu prov.	32.1	118.7	Eastern Wu	AD220-265	72.1	1.45	3
21. Brick	Gai county, Jilin prov.	40.4	122.4	Dong han (AD 141)	AD141	69.0	1.32	3
22. Sherd	Xian region, Shanxi prov.	34.4	109.2	Han	206BC-AD24	59.1	1.22	3
23. Sherd	Luoyang region, Henan prov.	34.2	112.7	Western Han	206BC-AD24	42.0	0.90	3
24. Brick	Luoyang region, Henan prov.	34.7	112.7	Han	206BC-AD24	55.3	1.14	2
25. Sherd	Guangzhou region, Guangdong prov.	23.0	113.1	Early Western Han	206BC-100BC	34.1	0.83	2
26. Sherd	Gui county, Guangxi prov.	23.2	109.7	Early Western Han	206-100BC	36.8	0.86	3
27. Brick	Luoyang region, Henan prov.	34.7	112.7	Western Han	206-100BC	45.3	0.94	3
28. Tile	Beijing region	40.0	116.2	Qin	221-207BC	54.5	1.06	3
29. Brick	Xian region, Shanxi prov.	34.4	109.2	Qin	221-207BC	45.2	0.93	1
30. Sherd	Xian region, Shanxi prov.	34.4	109.2	Qin	221-207BC	38.4	0.82	3
31. Sherd	Chenxi region, Hunan prov.	28.0	110.1	Late Warring States	351-221BC	42.5	0.99	3
32. Sherd	Chengdu region, Sichuan prov.	30.7	104.0	Warring States	351-221BC	43.0	0.97	3
33. Sherd	Zhengzhou region, Henan prov.	34.7	113.7	Warring States	475-350BC	47.8	1.02	3
34. Sherd	Luoyang region, Henan prov.	34.7	112.7	Warring States	475-350BC	41.7	0.87	2
35. Sherd	Zhengzhou region, Henan prov.	34.7	113.7	Warring States	770-476BC	48.2	1.03	3
36. Sherd	Zhengzhou region, Henan prov.	34.7	113.7	Warring States	770-476BC	45.8	0.99	3
37. Sherd	Zhengzhou region, Henan prov.	34.7	113.7	Warring States	770-476BC	51.5	1.07	3
38. Sherd	Zhengzhou region, Henan prov.	34.7	113.7	Warring States	770-476BC	57.0	1.18	2

continued overleaf

Table 1—continued

Sample	Site	Latitude (deg.)	Longitude (deg.)	Dynasty	Date	F^A (μT)	Intensity ratio	Grade
39. Sherd	Gaochun county, Jiangsu prov.	31.3	118.9	Warring States	570–476BC	57.4	1.23	3
40. Sherd	Jiangning county, Nanjing, Jiangsu prov.	32.1	118.7	Mid-Western Zhou	970–870BC	63.6	1.37	2
41. Sherd	Jiangning county, Nanjing, Jiangsu prov.	32.1	118.7	Late Shang	12th–11th Century BC	61.6	1.32	2
42. Sherd	Raoping county, Guangdong prov.	23.4	117.1	Shang	16th–11th Century BC	38.1	0.89	3
43. Sherd	Wu city, Jiangxi prov.	29.2	116.0	Shang	16th–11th Century BC	51.4	1.16	3
44. Sherd	Luoyang region, Henan prov.	34.7	112.7	Shang	16th–11th Century BC	67.8	1.40	3
45. Sherd	Luoyang region, Henan prov.	34.7	112.7	Shang	16th–11th Century BC	72.1	1.49	2
46. Sherd	Mengyang, Henan prov.	34.7	113.7	Mid-Shang	15th–14th Century BC	59.9	1.24	3
47. Sherd	Zhengzhou region, Henan prov.	34.7	113.6	Early Shang	16th–15th Century BC	65.2	1.35	3
48. Sherd	Mengyang, Henan prov.	34.7	113.7	Early Shang	16th–15th Century BC	58.0	1.20	3
49. Sherd	Luoyang region, Henan prov.	34.7	112.7	Late Xia	17th–16th Century BC	62.2	1.29	3
50. Sherd	Luoyang region, Henan prov.	34.7	112.7	Early Shang	16th–15th Century BC	49.2	1.05	3
51. Sherd	Mengyang, Henan prov.	34.7	113.7	Neolithic	~1800BC	41.4	0.86	3
52. Sherd	Zhechuan, Henan prov.	33.0	111.4	Neolithic	~2000BC	45.6	0.94	3
53. Sherd	Zhechuan, Henan prov.	33.0	111.4	Neolithic	~2500BC	49.7	1.03	3
54. Sherd	Zhechuan, Henan prov.	33.0	111.4	Neolithic	~2500BC	40.4	0.84	3
55. Sherd	Zhechuan, Henan prov.	33.0	111.4	Neolithic	~2500BC	40.2	0.83	3
56. Sherd	Dan city, Henan prov.	33.7	115.2	Neolithic	~3000BC	49.1	1.02	3
57. Sherd	Anxiang county, Hunan prov.	29.3	112.3	Neolithic	~5000BC	44.0	0.99	3
58. Sherd	Mengyang, Henan prov.	34.7	113.7	Neolithic	~3500BC	45.7	0.95	3
59. Sherd	Mengyang, Henan prov.	34.7	113.7	Neolithic	~3500BC	37.0	0.77	3
60. Sherd	Mengyang, Henan prov.	34.7	113.7	Neolithic	~3500BC	45.1	0.93	3
61. Sherd	Mengyang, Henan prov.	34.7	113.7	Neolithic	~3500BC	48.4	1.00	3
62. Sherd	Luoyang region, Henan prov.	34.7	112.7	Neolithic	~3700BC	45.4	0.94	3
63. Sherd	Mengyang, Henan prov.	34.7	113.7	Neolithic	~3800BC	38.0	0.79	3
64. Sherd	Lintong, Shanxi prov.	34.4	109.2	Neolithic	~3800BC	37.0	0.77	1
65. Sherd	Lintong, Shanxi prov.	34.4	109.2	Neolithic	~3800BC	36.3	0.75	3
66. Sherd	Mengyang, Henan	34.7	113.7	Neolithic	~4000BC	43.4	0.90	2
67. Sherd	Mengyang, Henan prov.	34.7	113.7	Neolithic	~4000BC	46.1	0.95	3
68. Sherd	Mengyang, Henan	34.7	113.7	Neolithic	~4000BC	42.0	0.87	3
69. Sherd	Mengyang, Henan prov.	34.7	113.7	Neolithic	~4000BC	39.8	0.82	3
70. Sherd	Mengyang, Henan prov.	34.7	113.7	Neolithic	~4000BC	39.6	0.82	3

The intensity ratio is F^A/F^D , where F^A is the ancient field intensity and F^D is the field intensity at the sample locality due to an axial geocentric dipole of present-day strength (8×10^{22} Am²). 'Grade' indicates reliability on the basis of the magnetic characteristics of the sample: 1, very good; 2, good; 3, poor.

and V. Jones for help in measuring many of the samples, and C. Shaw and A. Allsop for patient technical expertise.

Received 9 March; accepted 28 May 1987.

1. Smith, P. J. *Geophys. J. R. astr. Soc.* **12**, 213–362 (1967).
2. Cox, A. J. *geophys. Res.* **73**, 3257–60 (1968).
3. McElhinny, M. W. & Senanayake, W. E. *J. Geomagn. Geoelectr.* **34**, 39–51 (1982).
4. Aitken, M. J., Allsop, A. L., Bussell, G. D. & Winter, M. B. *Nature* **310**, 306–308 (1984).
5. Aitken, M. J., Alcock, P. A., Bussell, G. D. & Shaw, C. J. in *Geomagnetism of Baked Clays and Recent Sediments* (eds Creer, K. M., Tucholka, P. & Barton C. E.) 122–127 (Elsevier, Amsterdam 1983).
6. Stuiver, M. & Pearson, G. W. *Radiocarbon* **28**, 805–838 (1986).
7. Pearson, G. W. & Stuiver, M. *Radiocarbon* **28**, 839–862 (1986).
8. Beer, J. *et al.* *Nucl. Instrum. Meth.* **B5**, 380–384 (1984).
9. Thellier, E. & Thellier, O. *Ann. Geophys.* **15**, 285–367 (1959).
10. Aitken, M. J., Alcock, P. A., Bussell, G. D. & Shaw, C. J. *Archaeometry* **23**, 53–64 (1981).
11. Aitken, M. J., Allsop, A. L., Bussell, G. D. & Winter, M. J. *Geomagn. Geoelectr.* **38**, 1355–1363 (1986).
12. Barton, C. E., Merrill, R. T. & Barbetti, M. *Phys. Earth planet. Inter.* **20**, 96–110 (1979).
13. Bullard, E. C., Freedman, C., Gellman, H. & Nixon, J. *Phil. Trans. R. Soc.* **243**, 67–92 (1950).

Hominoid nasal region polymorphism and its phylogenetic significance

Robert B. Eckhardt

Department of Anthropology and Graduate Program in Genetics, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

The morphology of the nasal bones and their articulations with the adjoining frontal and maxillary bones have recently been reported in *Nature* and elsewhere^{1–3} to be diagnostic of hominoid taxa, and cladistic analysis based on these features has been used to assign two immature Plio-Pleistocene hominoids (AL 333-105 and Taung) to different lineages (*Paranthropus* and *Homo*, respectively). Because earlier studies^{4–10} had established that hominoid crania are highly variable intraspecifically, it seemed desirable to try to replicate the study reporting a consistently different pattern for each hominoid taxon³. My results show extensive polymorphism in the nasal region within every taxon of extant pongids, including several rather clear examples, in extant *Pan troglodytes*, of that recently hypothesized to be a 'paranthropine' pattern³. These findings underscore the substantial risks inherent in cladistic analyses using very restricted character sets to assign individual specimens to particular taxa.

Nasal bone outlines and the relationships among the frontal, nasal, and maxillary bones are said to be highly consistent within and between the hominoid taxa *Gorilla gorilla*, *Pan paniscus*, *Pan troglodytes*, *Pongo pygmaeus* and *Homo sapiens*³. Reportedly in great apes of both sexes and all age groups, nasion and glabella are closely approximated on the supraorbital torus; the nasal bones are widest immediately superior to the piriform aperture and converge superiorly; and the frontomaxillary sutures form a roughly transverse line below the level of nasion. The frontonasal and frontomaxillary sutures of *Homo sapiens*, however, are described as linear and continuous below an anteriorly projecting supraorbital torus, with nasion distant from the supraorbital torus and glabella.

The relevance of nasal morphology to early hominid systematics was reportedly enhanced by the discovery in robust australopithecines (also known as *Paranthropus*) of a third derived pattern. The lateral margins of the nasal bones are said to diverge superiorly, producing a 'keystone'-shaped frontal profile that more evenly distributes anterior occlusal forces through the supraorbital torus to the cranial vault.

Against the background of the earlier studies on hominoid primates^{4–10}, such a consistency of hominoid cranial suture patterning would not be expected. My investigation used 66 crania of *Pan troglodytes*, 99 of *Gorilla* and 108 of *Pongo* from the Mammal Collection at the U.S. National Museum of Natural

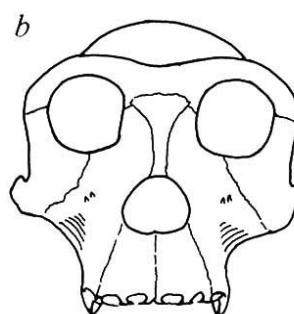
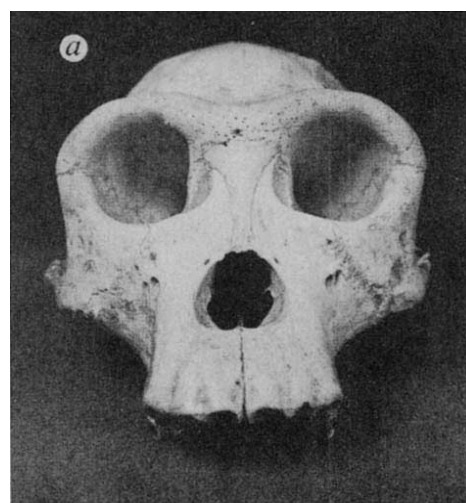


Fig. 1 *Pan troglodytes* (sex unspecified) with superiorly divergent nasomaxillary sutures forming keystone pattern said to be diagnostic of paranthropines.

History in Washington, D.C. and the Mammalogy Department at the Field Museum of Natural History in Chicago. In all three taxa both sexes were represented, in a wide range of ages from infancy to advanced adulthood. Most specimens had been collected in the wild.

My observations differed from those recently reported³ in several important respects: (1) my samples manifested substantial within-taxon variation in this anatomical region, particularly in the configuration of the nasomaxillary sutures; (2) there appear to be complex developmental changes affecting the relationships among the nasal, maxillary, and frontal bones through ontogeny; (3) the 'keystone' nasal bone arrangement suggested as a derived pattern diagnostic of *Paranthropus* is found in an appreciable number of pongids, particularly clearly in some chimpanzees.

The last point in the preceding list is the easiest to document unambiguously. Figure 1 shows a skull of *Pan troglodytes* (U.S. National Museum of Natural History 176238). This individual deviates substantially from the postulated sympleiomorphic pongid pattern, resembling, instead, the keystone pattern held to be characteristic of robust australopithecines. Moreover, USNMNH 176238 is not an isolated exception to these generalizations about the consistency of hominoid primate suture patterns and their phylogenetic significance.

In my own survey of 273 skulls of recent pongid primates, 23 specimens (8.4%) showed notable deviations from the patterns that the report characterized as consistent within and between taxa—that is, about one skull out of every twelve manifests some substantial exception. Tabulated separately by taxa, notable variants occurred in 9 of 99 *Pan*, 11 of 66 *Gorilla* and 3 of 108 *Pongo*. If anything, this attempt to quantify the extent of the exceptions to the postulated taxon-specific patterns

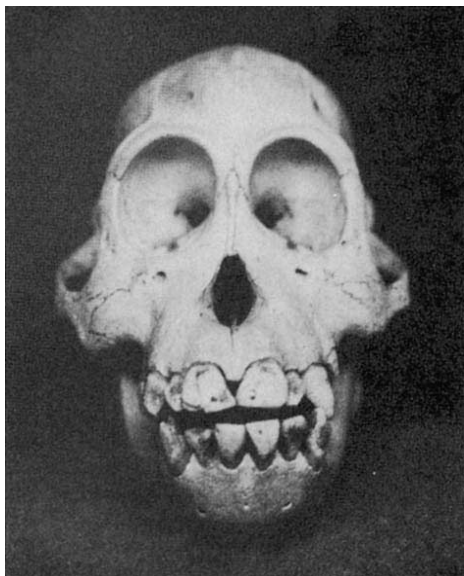


Fig. 2 *Pongo pygmaeus* (immature female) with nasal bones that are parallel to slightly divergent superiorly.

understates the variation in this anatomical region. A high proportion of the variants tabulated in my survey were specimens in which the nasal bones attained their greatest breadth well above the superior margin of the piriform aperture (Fig. 2). In some cases the nasals tapered to a blunt point near the piriform aperture instead of flaring widely (Fig. 3). In other instances, regardless of the outlines of the nasal bones, the sutures were deeply invaginated, even to the extent of containing relatively large accessory suture bones. The configurations observed were so multifarious (Fig. 4) that any attempt to designate one particular pattern as representative—as a character state—obscures the extensive structural polymorphism in this region.

Both specific and general conclusions follow from the evidence presented here. Because of its superiorly divergent nasal bone profile, in the recent report AL 333-105 from the Hadar Formation of Ethiopia was considered by Olson to exhibit a paranthropine pattern. But by Olson's own criteria the *Pan troglodytes* skull (USNMNH 176238) figured here manifests a

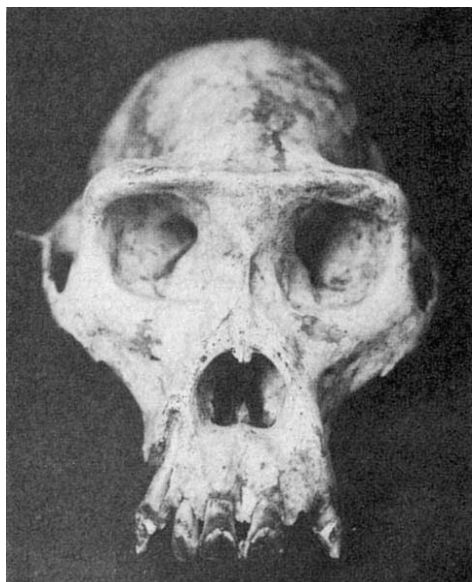


Fig. 3 *Gorilla gorilla* (female) with nasal bones that taper inferiorly to blunt point projecting into the piriform aperture.

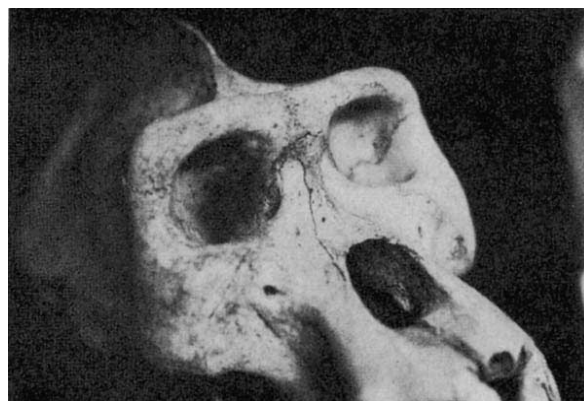


Fig. 4 *Gorilla gorilla* (male) with extremely complex sutures at junctures of nasal, maxillary, and frontal bones.

keystone pattern that is at least as 'paranthropine' as observed in any of the robust hominids described in that report. Consequently, the nasal morphology of AL 333-105 lends no support to the suggestion that there is taxonomic heterogeneity among the hominid sample from the Hadar Formation. More broadly, the occurrence of keystone patterns and other variants in *Pan*, *Gorilla* and *Pongo* populations vitiates the value of nasal bone morphology as an indicator of taxonomic affinity.

Nasal bone outlines are not only variable within a given hominoid taxon; sometimes they also exhibit similarities across quite different hominoid taxa. For example, an immature male *Pan troglodytes* (USNMNH 174703) has nasal bones that are narrower just above the upper rim of the piriform aperture and are more broadly-expanded, with rounded superior margins, near the glabellar region. Considered in isolation, the nasal bones of this specimen bear some resemblance to published photographs of KNM-WT 17000, a recently-discovered robust australopithecine¹¹, even though the rest of the cranial morphology differs strikingly between the two specimens. To trace the nasal bone outlines in KNM-WT 17000 with certainty, Walker and Leakey¹² found it desirable to use a scanning electron microscope; this finding illustrates some of the technical difficulties inherent in studying this anatomical region.

In hominoid primates the anatomical region including the nasal bones is so highly variable that to abstract a few patterns is seriously to misrepresent reality. Even if we accept a simple classificatory approach and tabulate the distribution of an idealized paranthropine keystone pattern, we find this 'paranthropine' pattern in over 8% of pongid specimens. Such a distribution suggests the existence of a polymorphism that is part of the common heritage of hominoid primates—not a character state limited to and diagnostic of any particular taxon. All of these findings call into question the earlier suggestion that nasal morphology provides a reliable basis for determining the phylogenetic affinities of individual adult or immature fossil crania. Single specimens unavoidably must fail to represent polymorphisms, whether these are under monogenic or polygenic influence. In attempting to reconstruct phylogenies, we must concern ourselves with the relationships among populations, not individuals.

I thank the U.S. National Museum of Natural History and the Field Museum of Natural History for hospitality, Henry Herpending and George Milner for helpful discussions, Frank B. Livingstone for drawing my attention to this problem.

Received 17 June 1986; accepted 28 May 1987.

1. Olson, T. R. *J. Hum. Evol.* 7, 159-178 (1978).
2. Olson, T. R. in *Ancestors: The Hard Evidence* (ed. Delson, E.) 102-119 (Liss, New York, 1985).
3. Olson, T. R. *Nature* 316, 539-540 (1985).
4. Coolidge, H. J. *Mem. Mus. Comp. Zool., Harvard* 1, 293-381 (1929).
5. Randall, F. E. *Hum. Biol.* 15, 236-254 (1943).

6. Randall, F. E. *Hum. Biol.* **15**, 307-337 (1943).
7. Randall, F. E. *Hum. Biol.* **16**, 23-76 (1944).
8. Schultz, A. H. *Contrib. Embryol.* **170**, Carnegie Inst. Washington Publ. **518**, 1-63 (1940).
9. Schultz, A. H. *Contrib. Embryol.* **182**, Carnegie Inst. Washington Publ. **525**, 57-110 (1941).
10. Tobias, P. V. *The Brain in Hominoid Evolution* (Columbia, New York, 1971).
11. Walker, A., Leakey, R. E., Harris, J. M. & Brown, F. H. *Nature* **322**, 517-522 (1986).
12. Walker, A. & Leakey, R. E. *Am. J. phys. Anthropol.* (in the press).

Seasonal and interannual changes in planktonic foraminiferal production in the North Pacific

Robert C. Thunell* & Susumu Honjo†

* Department of Geology, University of South Carolina, Columbia, South Carolina 29208, USA

† Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

It is well established that there are distinct seasonal cycles in the production of plankton, that vary both geographically and in terms of phytoplankton-zooplankton coupling¹. Only recently have attempts been made to measure seasonal variability in the flux of various biogenic components from the ocean surface to the seafloor²⁻⁴ and to relate this flux to local hydrographic conditions⁵. For example, a simple inverse relationship has been found between sediment flux to the seafloor and sea-surface temperature in the Sargasso Sea⁵. A three-year sediment trapping programme was begun at Ocean Station P (50° N, 145° W) in the subarctic North Pacific⁶. Two-week long flux samples have been collected continuously with an automated sediment trap²² at a depth of 3,800 m. The only gap in sample collection is from August 1984 to December 1984. For each of the samples collected we have determined the foraminiferal flux, as both mass and number of shells, for the 63 µm to 1 mm size fraction. In addition, total carbonate flux was measured for three discrete size fractions (<63 µm, 63 µm-1 mm, >1 mm).

In September 1982 a long-term sediment trapping programme was begun at Ocean Station P (50° N, 145° W) in the subarctic North Pacific⁶. Two-week long flux samples have been collected continuously with an automated sediment trap²² at a depth of 3,800 m. The only gap in sample collection is from August 1984 to December 1984. For each of the samples collected we have determined the foraminiferal flux, as both mass and number of shells, for the 63 µm to 1 mm size fraction. In addition, total carbonate flux was measured for three discrete size fractions (<63 µm, 63 µm-1 mm, >1 mm).

Seasonal fluctuations in the latitude position of the North Pacific Drift (NPD) strongly influence the thermal structure of the upper part of the water column in the region of Ocean Station P (ref. 7). During the summer, when the NPD is just south of Ocean Station P, sea-surface temperatures reach their annual maximum of ~14 °C, in September. At this time of year the near-surface layer is well stratified because of a strong thermocline. In contrast, the surface layer becomes well mixed during the winter months as the NPD migrates to the south. Annual surface water temperature cycles for Ocean Station P for January 1982 through June 1985 are shown in Fig. 1. For most of the study period (September 1982-April 1985) sea-surface temperatures at this location were slightly warmer than normal, in part because of the 1982-1983 El Niño event^{8,9}.

The subarctic North Pacific is an unusual pelagic ecosystem in that it is not characterized by any significant seasonal cycle of phytoplankton abundance¹⁰. At Ocean Station P, the standing stock of plants in the surface mixed layer, measured in terms of chlorophyll *a* concentration, averages ~0.3 mg m⁻³ throughout the year¹⁰. This is in sharp contrast to the North Atlantic where phytoplankton blooms occur in response to the seasonal development and decay of the thermocline.

The time series records for Ocean Station P clearly indicate that there are significant seasonal and interannual differences in foraminiferal flux (Fig. 2a). The planktonic foraminiferal

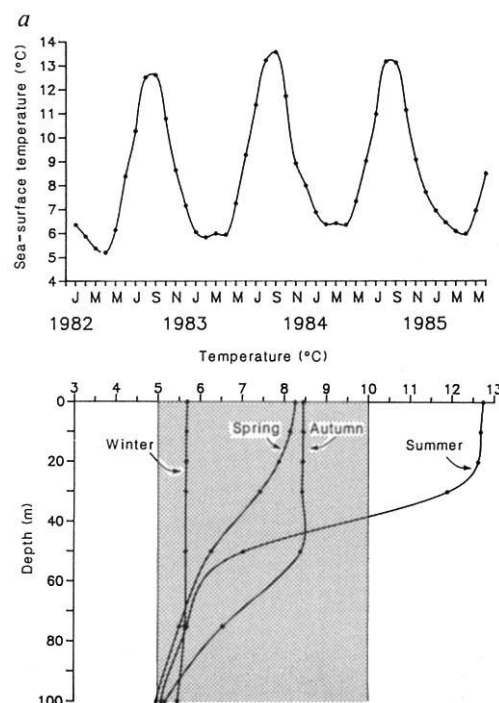


Fig. 1 *a*, Monthly mean sea-surface temperatures for Ocean Station P for January 1982 to May 1985. *b*, Typical seasonal temperature profiles for Ocean Station P. The shaded region represents the 5-10 °C temperature range, the optimal range for sub-polar planktonic foraminifera.

results show that there are two periods of high flux during the year, one in the spring (April-May) and one in the autumn (October-November), and that the magnitudes of these peaks vary considerably from year to year (Fig. 2a). From the available data it is not readily apparent if either the spring or autumn peak in foraminiferal flux is consistently larger than the other.

All three yearly records show that foraminiferal flux is consistently low during the winter (December-March) (Fig. 2a). The summer flux pattern is more difficult to resolve. Foraminiferal flux decreases in June of each year. In 1984, the flux remained low throughout the summer, whereas in 1983 and 1985 there were brief increases in foraminiferal flux during July followed by decreases in August and September.

The carbonate flux record (Fig. 2) includes both aragonite, primarily pteropods, and calcite, predominantly coccolithophorids and planktonic foraminifera. The coccoliths make up most of the fine-fraction calcite (<63 µm) and the foraminifera are most abundant in the coarse fraction (>63 µm). The total carbonate flux record (Fig. 2b) reveals a very distinctive seasonal cycle that is similar to that observed for foraminiferal flux. Carbonate flux is high during the autumn and spring, and low during the winter. It also appears that the spring peak is larger in magnitude but shorter in duration than the autumn peak (Fig. 2b). The summer records are somewhat inconsistent; carbonate flux is very high during July 1983 and low during July 1984 and 1985. We are assuming that the summer flux records for 1984 and 1985 are the norm and that the high July 1983 record is anomalous. This region was still being influenced by El Niño during the summer of 1983; the anomalously high flux during July 1983 may somehow be related to this event. Unusually high fluxes of carbonate during the summer of 1983 in the north equatorial Pacific have also been attributed to changes in surface water productivity induced by El Niño (J. Dymond and R. Collier, unpublished results).

The observed foraminiferal and carbonate flux patterns seem to contradict what is known about plankton ecology and production at this location in the subpolar North Pacific (Fig. 3). Phytoplankton standing stock is relatively uniform throughout

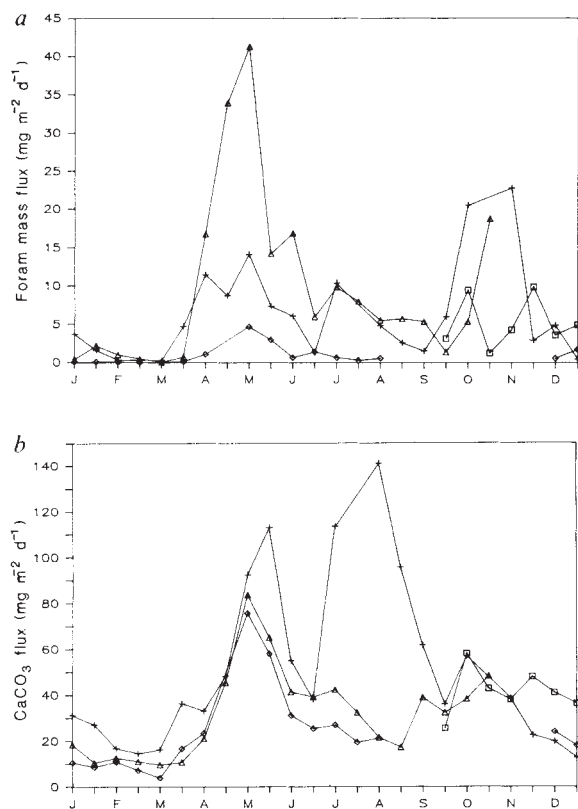


Fig. 2 Monthly foraminiferal and carbonate flux data for: $\square$, 1982; $+$, 1983; $\diamond$, 1984; $\triangle$, 1985. **a**, Planktonic foraminiferal flux estimates for time-series sediment trap samples collected at Ocean Station P. **b**, Carbonate flux estimates for time-series sediment trap samples collected at Ocean Station P.

the year at Ocean Station P^{10,11,12}. However, there is a very distinctive seasonal cycle of primary production (Fig. 3a) which is controlled by the annual cycle of incident solar radiation^{13,14} as reflected in the annual sea-surface temperature pattern (Fig. 1a). The annual cycle of zooplankton standing stock (Fig. 3b) mirrors that of primary production (Fig. 3a), with each reaching a maximum in mid-summer^{10,15}. Therefore an increase in zooplankton abundance is not dependent upon a phytoplankton bloom as it is in the North Atlantic. In addition, zooplankton grazing must maintain the relatively uniform phytoplankton abundances found throughout the year. How do we reconcile and interpret our flux results in light of these established plankton relationships? The autumn and spring increases in foraminiferal and carbonate flux at Ocean Station P would seem much more characteristic of a North Atlantic-like setting⁵.

Individual species of planktonic foraminifera have specific temperature tolerance ranges¹⁶ and thermal conditions in the photic zone control both the vertical distribution and seasonal succession of these zooplankton^{17,18}. Seasonal changes in the thermal structure of the upper water column will result in changes in both the thickness and depth of the preferred temperature habitat of individual planktonic foraminiferal species. Food availability is also an important factor in regulating foraminiferal abundance in the water column¹⁷. Planktonic foraminifera are passively drifting organisms, utilizing whatever food they encounter¹⁹. In particular, marine copepods are a major food source for planktonic foraminifera¹⁹ and in our study area three species of copepods account for 90% of the zooplankton biomass¹⁰. Many of the spinose planktonic foraminiferal species also possess symbiotic algae which provide additional nutrition for their hosts²⁰.

The relationship in time between surface layer thermal structure and food availability may be the key to the observed

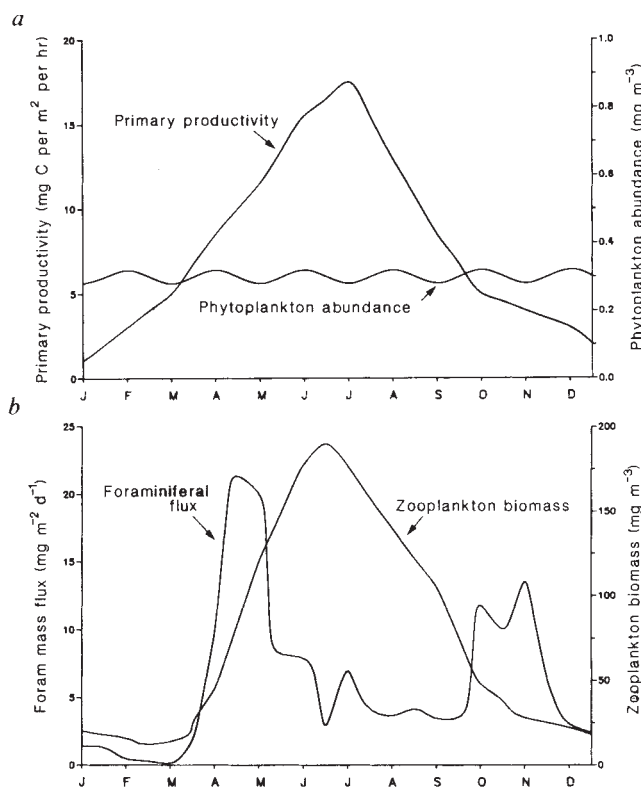


Fig. 3 **a**, Mean seasonal primary productivity record for the upper 50 m at Ocean Station P (based on data from ref. 13). Mean phytoplankton abundance record for Ocean Station P (based on data from ref. 11), averaged over the years 1971–1976. **b**, Mean zooplankton biomass record for Ocean Station P (based on data from ref. 15). Mean foraminiferal flux record for Ocean Station P derived from the individual year-long records of Fig. 2.

foraminiferal flux patterns (Fig. 2a). It is reasonable to assume that there are certain surface layer temperature and food source associations that are optimal for foraminiferal production, whereas other associations would result in low foraminiferal productivity. For example, we expect maximum production at times when food availability is high and thermal structure is optimal for a specific foraminiferal assemblage. Likewise, production should be low during periods when temperatures are either above or below the tolerance range of an assemblage or the thickness of the preferred habitat is greatly reduced, even if food availability is high.

At Ocean Station P we are dealing almost exclusively with a subpolar assemblage consisting of *Globigerina bulloides*, *Globigerina quinqueloba*, *Globigerinita glutinata*, *Globorotalia scitula* and *Neoglobobulimina pachyderma*²¹. The optimal temperature range for this assemblage is 5–10 °C (ref. 16) and it undergoes a very distinctive pattern of seasonal succession²¹. The only non-subpolar species collected in the sediment traps is *Orbulina universa*. This species is more characteristic of subtropical to transitional regions¹⁶, and only occurs at Ocean Station P during the summer when sea-surface temperatures are highest.

By considering the seasonal changes in surface layer thermal structure and food availability, together with the temperature tolerances of the planktonic foraminifera present at Ocean Station P, we can establish a model that explains the observed foraminiferal flux patterns. During the winter, primary productivity is low owing to low light intensity, and the standing stock of zooplankton (primarily copepods) is also low (Fig. 3b). During this period the surface layer is well mixed (Fig. 1a) and temperatures throughout this layer are close to the lower limit of the tolerance range for subpolar planktonic foraminiferal species. The combination of low food availability and low tem-

peratures results in low foraminiferal production and flux during the winter.

A simultaneous increase in primary productivity and zooplankton abundance occurs during the spring (Fig. 3), along with the development of a thermocline (Fig. 1). The increase in foraminiferal food source, both in terms of copepod abundance and necessary light intensity levels for symbiotic algae, together with a thick 5–10 °C habitat (surface to 100 m depth) results in increased production and flux of planktonic foraminifera during the spring.

Although food availability remains high, there is a sharp decrease in foraminiferal flux during most of the summer—probably because the surface layer becomes very stratified and temperatures in the upper 40 m exceed the tolerance range for subpolar species (Fig. 1). During the summer, the thickness of the 5–10 °C habitat is reduced and is restricted to the lower half of the thermocline. As previously stated, *O. universa* is the only non-subpolar species present at Ocean Station P²¹, and it accounts for much of the summer flux. This subtropical-transitional species most probably lives in the surface mixed layer during this period. Zooplankton grazing may also contribute to the low foraminiferal fluxes during the summer. Zooplankton abundance is greatest at this time of year, so grazing pressure should also be at a maximum, but the data required to assess the impact of grazing on foraminiferal flux are not available.

During the autumn, the thermocline begins to decay and mixed layer temperatures decline back into the optimal range for subpolar species. This results in a vertical expansion of the 5–10 °C habitat for these species. These changes in thermal structure combined with a relatively high food availability result in an increase in foraminiferal production and flux.

Significantly more interannual variability exists in the foraminiferal flux record than in the carbonate flux record (Fig. 2a and b). In fact, the carbonate flux values are remarkably similar from year to year. The greatest amount of interannual variability in foraminiferal flux is associated with the spring and autumn periods of high flux. The variability in foraminiferal flux does not appear to be directly related to anomalous sea-surface temperatures (Fig. 1a). For example, the spring flux peaks range from a low of 5 mg m⁻² d⁻¹ in May 1984 to a high of 40 mg m⁻² d⁻¹ in May 1985. However, sea-surface temperatures in May of all three years differ by less than 0.4 °C. Year-to-year differences in the sizes of the flux peaks may be a function of the level of food availability. Unfortunately, the data necessary to test this possibility are not available.

We thank C. S. Wong (Institute of Ocean Sciences, British Columbia) for help throughout the sediment trapping project at Ocean Station P, R. Fairbanks, W. B. Curry and K. Takahashi for helpful comments on the manuscript, and S. Manganini, R. Bellegay and E. Tappa for technical and analytical assistance. This research was supported by NSF grants to R.T. and S.H.

Received 5 March; accepted 12 May 1987.

- Heinrich, A. K. *J. Cons. perm. int. Explor. Mer.* **27**, 15–24 (1962).
- Deuser, W. G. & Ross, E. H. *Nature* **283**, 364–365 (1980).
- Honjo, S. *Global Ocean Flux Study*, 305–324 (National Academy of Sciences, Washington, DC, 1984).
- Honjo, S. *Science* **218**, 883–884 (1982).
- Deuser, W. G. *Deep Sea Res.* **33**, 225–246 (1986).
- Honjo, S. & Wong, C. S. *EOS* **66**, 1282 (1985).
- Livingston, D. & Royer, T. C. *J. phys. Oceanogr.* **10**, 753–764 (1980).
- Royer, T. C. & Xiong, Q. *Top. Ocean-Atmos. Newsl.* **24**, 4–5 (1984).
- Reed, R. K. *Mar. Fish. Rev.* **46**, 7–12 (1984).
- Miller, C. B., Frost, B. W., Batchelder, H. P., Clemons, M. J. & Conway, R. E. *Prog. Oceanogr.* **13**, 201–243 (1984).
- Anderson, G., Lam, R., Booth, B. & Glass, J. *Spec. Rep. Dep. Oceanogr. Univ. Wash.* **76**, 1–231 (1977).
- Frost, B. W., Landry, M. R. & Hassett, R. P. *Deep Sea Res.* **30**, 1–13 (1983).
- Stephens, K. *Fish. Res. Board Can. Manuscr. Rep. Ser.* **257** (1968).
- Frost, B. W. *Global Ocean Flux Study*, 125–135 (National Academy of Sciences, Washington DC, 1984).
- Fulton, J. D. *Can. Date Rep. Fish. Aquat. Sci.* **374**, 1–65 (1983).
- Bé, A. W. H. *Oceanic Micropaleontology* Vol. 1 (ed. Ramsay, A.) 1–100 (Academic Press, 1977).
- Fairbanks, R. G. & Wiebe, P. H. *Science* **209**, 1524–1526 (1980).
- Fairbanks, R. G., Sverdrup, M., Free, R., Wiebe, P. H. & Bé, A. W. H. *Nature* **298**, 841–844 (1982).

- Bé, A. W. H. in *Foraminifera: Notes for a Short Course* Vol. 6 (ed. Broadhead, T. W.) 51–92 (Univ. Tennessee, 1982).
- Hemleben, C. & Spindler, M. *Utrecht Micropaleont. Bull.* **30**, 141–171 (1983).
- Reynolds, L. & Thunell, R. C. *J. foram. Res.* **15**, 282–301 (1985).
- Honjo, S. & Doherty, K. W. *Deep Sea Res.* (in the press).

Rhizobium nodulation gene *nodD* as a determinant of host specificity

Herman P. Spaiak, Carel A. Wijffelman, Elly Pees, Rob J. H. Okker & B. J. J. Lugtenberg

Department of Plant Molecular Biology, Leiden University, Nonnensteeg 3, 2311 VJ Leiden, The Netherlands

The symbiosis between bacteria of the genus *Rhizobium* and their leguminous host plants results in the formation of root nodules in a species-specific way, in that a particular bacterial species can nodulate only a limited number of host species. In fast-growing *Rhizobium* species many nodulation (*nod*) genes, including the functional interchangeable or common *nod* genes *nodA, B, C, I, J* and *nodD* and the host-specificity genes *nodE, F*, are localized on large Sym (for symbiosis) plasmids^{1–5}. On the *Rhizobium meliloti* Sym plasmid three *nodD* genes have been localized. Two of these, *nodD1* and *nodD2*, are required for efficient nodulation of the host plant (ref 6 and M. Honma, personal communication). Exudates of leguminous plants induce the expression of several Sym-plasmid-localized *nod* operons^{7–10}, a process in which the constitutively expressed *nodD* product is supposed to act as a positive regulator^{8,9}. The inducing compounds found in these exudates have been identified as flavones, flavanones or closely related compounds^{11–14}. We report here that the *nodD* products of the various fast-growing *Rhizobium* species differ from each other in that they confer different responsiveness, in a species-specific way, to different sets of flavonoids and exudates. Moreover, in one case, the *nodD* gene was shown to be a determinant of host-specific nodulation.

The function of the *nodD* product in the recognition of the plant flavonoid inducer of the *nod* genes is unknown. It is striking that, although *nodD* is considered to be a common *nod* gene, the structural requirements for inducers of *nod* genes of *R. leguminosarum*^{13,14}, *R. trifolii*¹² and *R. meliloti*¹¹ seem to be different. Therefore, the *nodD* genes of the fast-growing *Rhizobium* species *R. leguminosarum*, *R. trifolii* and *R. meliloti* (the *nodD1* gene) were cloned (Fig. 1) in a broad host range IncP class vector and the resulting plasmids pMP280, pMP283 and pMP284 were subsequently transferred to *Rhizobium* strain LPR5045 which lacks a Sym plasmid. The inserts contain a complete *nodD* gene, including its own constitutively expressed promoter and some surrounding sequences, but no other complete *nod* gene^{2,7–9,15–18} (Fig. 1). To measure the expression of inducible *nod* promoters, plasmid pMP154 (Fig. 1) was used, an IncQ transcriptional fusion vector in which a 114-base pair (bp) fragment containing the upstream region (including the presumed promoter¹⁹) of the *nodA, B, C, I, J* operon of the *R. leguminosarum* Sym plasmid pRL1J1 was cloned. In pMP154, *nod* promoter activity could be detected as β -galactosidase activity. The upstream regions with inducible promoter activity of the *nodA, B, C* operons of *R. trifolii* and *R. meliloti* were cloned in the IncQ transcriptional fusion vector resulting in pMP193 and pMP194, respectively (Fig. 1) and promoter activity could be detected as with construct pMP154. Plasmid pMP154 was mobilized to the *Rhizobium* LPR5045 derivatives harbouring the *nodD*-containing plasmids, yielding an isogenic set of strains which differ only in the source of the *nodD* genes.

The results of the induction experiments (Table 1, lines 1–3) show that luteolin can induce the *nod* promoter to a substantial level regardless of the *nodD* source. For the other flavonoids the source of *nodD* determines which substances efficiently

Fig. 1 Cloning of *nodD* genes from various *Rhizobium* species and of *nodA,B,C* upstream regions. The *nodD* genes of the *R. leguminosarum* Sym plasmid pRL1JI (ref. 25), the *R. trifolii* Sym plasmid pRtr843 and the *R. meliloti* Sym plasmid pRmeSU47a, present in the plasmids pMP104 (ref. 5), pRt308 (ref. 4) and pRmeSL26 (ref. 30), respectively, were subcloned (hatched bars) in the vector pMP92 (ref. 5). This IncP vector is 7 kilobases (kb) in size, codes for tetracycline resistance and contains a polylinker. The resulting constructs pMP280, pMP283 and pMP284 were mobilized to the *Rhizobium* strains used in this study. A 114-basepair (bp) *Bgl*II-*Bcl*I fragment, containing inducible promoter activity and the *nod*-box (defined according to ref. 24) of the *nodA,B,C,I,J* operon of pRL1JI, was cloned in the transcriptional fusion vector pMP190 (ref. 5) with the indicated direction towards *lacZ*, resulting in the plasmid pMP154 (ref. 5). The indicated fragments, which contain inducible promoter activity of the *nodA,B,C* operons of pRtr843 and pRmeSU47a were also cloned in pMP190 and resulted in the constructs pMP193 and pMP194, respectively. Vector pMP190, of the IncQ class, codes for chloramphenicol and streptomycin resistance and contains the *Escherichia coli lacZ* gene, without its promoter and operator region, as an indicator gene for transcriptional expression. Part of the *nod* regions, the restriction sites used, the conserved *Bam*HI site, and the localization of the *nod* genes in these regions, are indicated. Size and position of the *nod* genes, indicated by solid lines, are according to refs 2 and 15-18 except that the *nodA* translational start of pRL1JI is drawn 91 bp closer to the *Bcl*I site (unpublished results) than indicated in other studies^{2,15}, which makes this *nodA* product more homologous to the *nodA* products of other fast-growing *Rhizobium* species. The parts of *nodC* and *nodE* of *R. trifolii*, for which nucleotide sequences have not yet been reported, are indicated tentatively by dotted lines. *Nod*-boxes, strongly conserved DNA sequences which precede every inducible *nod*-operon^{2,5,17,19,24,31}, are indicated by triangles. Restriction enzyme sites are shown as follows: B, *Bam*HI; Bc, *Bcl*I; Bg, *Bgl*II; C, *Cla*I; H, *Hind*III; K, *Kpn*I; Sp, *Sph*I; Ss, *Sst*I.

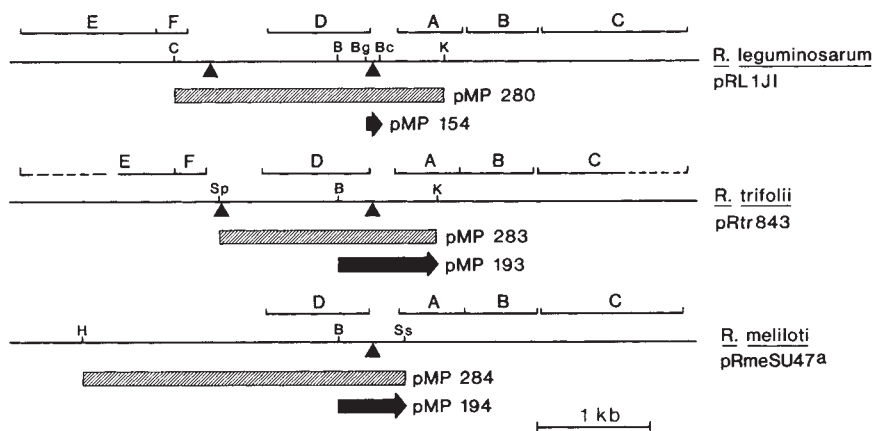


Table 1 Expression of *R. leguminosarum nodABCIIJ* promoter by commercially available flavonoids and natural inducers in the presence of *nodD* genes from various sources

<i>nodD</i> -containing plasmid	Units β -galactosidase ($\times 10^{-3}$)						<i>Trifolium pratense</i> exudate	<i>T. repens</i> exudate
	No inducer	Luteolin	Apigenin	Naringenin	Eriodictyol	7-Hydroxy-flavone		
pMP280 (<i>R. leguminosarum</i>)	0.26	19	24	21	21	7.6	0.41	14
pMP283 (<i>R. trifolii</i>)	0.84	22	29	28	4.6	29	8.0	30
pMP284 (<i>R. meliloti</i>)	0.30	6.0	2.6	0.54	0.90	0.55	0.40	7.5
pRL1JI (<i>R. leguminosarum</i>)	0.30	18	20	19	19	7.3	0.43	9.3
pSym1 (<i>R. leguminosarum</i>)	0.61	13	13	14	13	4.8	0.69	10
pPRE (<i>R. leguminosarum</i>)	0.43	14	14	16	12	5.2	0.62	14
pHIM (<i>R. leguminosarum</i>)	0.40	19	17	17	16	7.2	0.90	17
pRtr843 (<i>R. trifolii</i>)	0.90	24	29	26	4.8	29	9.2	24
pSym5 (<i>R. trifolii</i>)	1.9	14	19	17	6.3	10	10	15

Rhizobium derivatives that contain an IncP plasmid, with one of the *nodD* genes, as well as the IncQ plasmid pMP154 (Fig. 1) were constructed by introducing these plasmids into strain LPR5045 which has been cured for its own Sym plasmid²³. When intact Sym plasmids were to be used as the source of *nodD*, Tn5-marked Sym plasmids were transferred to LPR5045 either by direct conjugation (pRL1JI, pHIM and pSym5) or by mobilization with the helper plasmid R180 (ref. 23), and pMP154 was introduced subsequently. The Sym plasmids pRL1JI (ref. 25) pSym1 (ref. 23) and pPRE (ref. 26) confer nodulation ability on *Pisum sativum* cv Rondo, *Vicia sativa* ssp. nigra and *V. hirsuta*, whereas pHIM (ref. 26) confers the ability to nodulate these *R. leguminosarum* hosts as well as *P. sativum* cv Afghanistan (a Middle East variety). The Sym plasmids pRtr843 (ref. 4) and pSym5 (ref. 23) allow nodulation on *Trifolium pratense* and *T. repens*. *Nod* promoter induction assays were performed as described previously¹⁴ after incubation for 12 h (maximal induction), if applicable in the presence of 400 nM flavonoid or undiluted exudate, and the measured activities are indicated as units of β -galactosidase²⁷. The induction assays were performed in triplicate and variation from the given value was within 10%. Exudates of *Melilotus alba*, *P. sativum*, *V. hirsuta*, *T. repens* and *T. pratense* were prepared as described previously¹⁰. The former four exudates had similar characteristics in that they induced the *nod* promoter in each strain to at least 50% of the induction level observed with luteolin. Only representative results (with *T. repens* exudate) are shown. The following control experiments were performed. To test whether the absence of inducing capacity of *T. pratense* exudate with several strains was due to the presence of inhibitors, the level of induction by 400 nM luteolin was compared with that of a mixture of 400 nM luteolin and undiluted *T. pratense* exudate. Induction levels were found to be identical. *Rhizobium* LPR5045 containing pMP154 (with no *nodD* gene present) showed 150 units β -galactosidase with all tested compounds. Apigenin, eriodictyol, and luteolin were from Carl Roth, Karlsruhe. Naringenin was from Sigma and 7-hydroxyflavone from EGA, Steinheim, FRG.

Table 2 Nodulation of *T. pratense* and *T. repens* by *R. trifolii* derivatives containing *nodD* genes of various *Rhizobium* species

Rhizobium strain (and <i>nodD</i> source)	Nodulated plants (%)		Average number of nodules per nodulated plant	
	<i>T. pratense</i>	<i>T. repens</i>	<i>T. pratense</i>	<i>T. repens</i>
ANU851.pMP280 (<i>R. leguminosarum</i>)	3	100	1	4
ANU851.pMP283 (<i>R. trifolii</i>)	90	100	2	4
ANU851.pMP284 (<i>R. meliloti</i>)	0	70	—	2
ANU851 (<i>nodD</i> ::Tn5)	0	0	—	—
ANU843 (<i>R. trifolii</i> wild type)	90	100	2	4

The plasmids pMP280, pMP283 and pMP284 (Fig. 1) were mobilized to the *R. trifolii* strain ANU851, which contains a Tn5 insertion in the *nodD* of its Sym plasmid pRt843 (ref. 28). The resulting strains, as well as ANU851 and the wild-type strain ANU843, were tested for nodulation ability on *T. repens* (white clover) and *T. pratense* (red clover). For each strain, 60 plants were infected and nodulation assays were performed as described previously²⁹. All nodulated plants were able to fix nitrogen. Plant seeds were from Kieft, Blokker (Netherlands).

activate the *nod*-inducible promoter. The strains containing the *nodD* gene of *R. leguminosarum* or *R. trifolii* can be most clearly distinguished by their reactions with eriodictyol and 7-hydroxyflavone in that eriodictyol is the better inducer in the presence of *R. leguminosarum nodD* (ratio eriodictyol/7-hydroxyflavone is 2.8) whereas 7-hydroxyflavone is a much better inducer in the presence of *R. trifolii nodD* (ratio eriodictyol/7-hydroxyflavone is 0.16). In contrast, neither these two substances nor naringenin were efficient inducers for the strain containing the *R. meliloti nodD* gene (Table 1). In the case of *R. leguminosarum* and *R. trifolii* the same differences in induction were also observed when the complete Sym plasmids pRL1JI and pRt843 were used as the source of *nodD* (Table 1, lines 4 and 8). As a control, the strains containing the combinations pMP283 plus pMP193 and pMP284 plus pMP194, representing combinations of *nodD* clones with their own inducible promoters, were tested with the same inducers (Table 1). Almost similar results were obtained as with the inducible promoter in pMP154, although the combination pMP283 plus pMP154 showed a higher level of induction than the combination pMP283 plus pMP193 (data not shown).

To decide whether the observed differences in induction patterns could have species-specific basis, the Sym plasmids of three other *R. leguminosarum* strains and one other *R. trifolii* strain were tested as the *nodD* source (Table 1, lines 5, 6, 7 and 9). The three other *R. leguminosarum* Sym plasmids yielded essentially the same results as observed with pRL1JI as the *nodD* source. Similarly, the results obtained with the other *R. trifolii* Sym plasmid were comparable with that of pRt843. The data in Table 1 therefore indicate that the *nodD* products of the *Rhizobium* species tested differ in their response to flavonoid inducers in a species-specific way.

To see whether this specificity could also be reflected in nature, exudates of *Melilotus alba*, *Pisum sativum*, *Vicia hirsuta*, *Trifolium repens* and *T. pratense* were tested for their ability to induce the *nod* promoter of pMP154 in the presence of the various *nodD* clones. The former four exudates induced in each *nodD* strain the *nod* promoter to at least 50% of the induction level of luteolin (Table 1). In contrast, exudate of *T. pratense* (red clover) was only a good inducer in the presence of the *nodD* gene of *R. trifolii* (Table 1). We have also shown that the absence of induction with red clover exudate was not due to inhibitors (Table 1 legend). Because the control with *T. repens* exudate shows that induction with clover exudate is possible in all cases, these results confirm that the *nodD* genes of the tested plasmids have cross-inoculation group specific characteristics.

The results with the *T. pratense* exudate cast doubt on the general assumption that *nodD* is a common *nod* gene. We therefore investigated whether the *nodD* gene of either *R. leguminosarum* or *R. meliloti* can function in the nodulation of

red clover. Therefore the plasmids with the different *nodD* genes were mobilized into *R. trifolii* strain ANU851 (pRt843 *nodD*::Tn5). When the resulting strains were tested for nodulation ability on red and white clover, it appeared that the *nodD* genes of *R. leguminosarum* and *R. meliloti* are only able to complement the mutation of strain ANU851 for nodulation on white clover (and the *R. meliloti nodD* not completely) whereas the (control) *nodD* gene of *R. trifolii* can complement for nodulation on both *Trifolium* species (Table 2). We conclude that the host range of *R. trifolii* is narrowed when its own *nodD* gene is replaced by that of *R. leguminosarum* or *R. meliloti* and that therefore *nodD* cannot be considered as a common *nod* gene.

Although the *nodD* products of *R. leguminosarum*, *R. trifolii* and *R. meliloti* have a homology with each other of at least 75% (refs 2, 16 and 17), they are strikingly different in a number of properties. In an isogenic series of strains which differ only in the source of their *nodD* gene, it was found that these strains differ in their response to a set of flavonoid inducers (Table 1). For the *R. leguminosarum* and *R. trifolii* Sym plasmids tested, this response appears to be species specific (Table 1). In at least one case, namely the symbiosis between *R. trifolii* and *T. pratense*, the *nodD* gene does not behave as a common *nod* gene (Table 2). It has been reported that the *nodD* genes of fast-growing *Rhizobium* species can complement each other in nodulation assays on several legume host plants. Therefore *nodD* has been designated as a common *nod* gene^{1,20-22}. Our results show that this notion should be revised. The results indicate that the interaction between flavonoid inducer and *nodD* product is crucial in the nodulation process. To extend the host range of nodulation, even to non-leguminous plants, strategies based upon manipulation of the interaction between these molecules could be successful.

We thank Ine Mulders for technical assistance and Bas Zaat for unpublished information about exudates.

Received 25 February; accepted 3 June 1987.

- Kondorosi, E. & Kondorosi, A. *Trends biochem. Sci.* **11**, 296-299 (1986).
- Shearman, C. A., Rossen, L., Johnston, A. W. B. & Downie, J. A. *EMBO J.* **5**, 647-652 (1986).
- Evans, I. J. & Downie, J. A. *Gene* **43**, 95-101 (1986).
- Djordjevic, M. C., Innes, R. W., Wijffelman, C. A., Schofield, P. R. & Rolfe, B. G. *Pl. Molec. Biol.* **6**, 389-401 (1986).
- Spaink, H. P., Okker, R. J. H., Wijffelman, C. A., Pees, E. & Lugtenberg, B. J. J. in *Recognition in Microbe-Plant Symbiotic and Pathogenic Interactions* (ed. Lugtenberg, B. J. J.) 55-68 (Springer, Heidelberg, 1986).
- Göttfert, M. *et al. J. molec. Biol.* **191**, 411-420 (1986).
- Innes, R. W. *et al. Molec. gen. Genet.* **201**, 426-432 (1985).
- Mulligan, J. T. & Long, S. R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6609-6613 (1985).
- Rossen, L., Shearman, C. A., Johnston, A. W. B. & Downie, J. A. *EMBO J.* **4**, 3369-3373 (1985).
- Van Brussel, A. A. N. *et al. J. Bact.* **165**, 517-522 (1986).
- Peters, N. K., Frost, J. & Long, S. *Science* **233**, 977-980 (1986).
- Redmond, J. W. *et al. Nature* **323**, 632-635 (1986).
- Firmin, J. L., Wilson, K. E., Rossen, L. & Johnston, A. W. B. *Nature* **324**, 90-94 (1986).
- Zaat, S. A. J. *et al. J. Bact.* **169**, 198-204 (1987).
- Rossen, L., Johnston, A. W. B. & Downie, J. A. *Nucleic Acids Res.* **12**, 9497-9508 (1984).

16. Egelhoff, T. T., Fisher, R. F., Jacobs, T. W., Mulligan, J. T. & Long, S. R. *DNA* 4, 241-248 (1985).
17. Schofield, P. R. & Watson, J. M. *Nucleic Acids Res.* 14, 2891-2903 (1986).
18. Jacobs, T. W., Egelhoff, T. T. & Long, S. R. *J. Bact.* 162, 469-476 (1985).
19. Spink, H. P., Okker, R. J. H., Wijffelman, C. A., Pees, E. & Lugtenberg, B. *Pl. molec. Biol.* (in the press).
20. Djordjevic, M. A. *et al. Pl. molec. Biol.* 4, 147-160 (1985).
21. Wijffelman, C. A., Pees, E., Van Brussel, A. A. N., Okker, R. J. H. & Lugtenberg, B. J. J. *Arch. Microbiol.* 143, 225-232 (1985).
22. Fisher, R. F., Tu, J. K. & Long, S. R. *Appl. env. Microbiol.* 49, 1432-1435 (1985).
23. Hooykaas, P. J. J., Snijderwint, F. G. M. & Schilperoort, R. A. *Plasmid* 8, 73-82 (1982).
24. Rostas, K., Kondorosi, E., Horvath, B., Simoncsits, A. & Kondorosi, A. *Proc. natn. Acad. Sci. U.S.A.* 83, 1757-1761 (1986).
25. Johnston, A. W. B. *et al. Nature* 276, 635-636 (1978).
26. Winarno, R. & Lie, T. A. *Plant. Soil* 51, 135-142 (1979).
27. Miller, J. H. *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, New York, 1972).
28. Djordjevic, M. A., Schofield, P. R. & Rolfe, B. G. *Molec. gen. Genet.* 200, 463-471 (1985).
29. Van Brussel, A. A. N., Tak, T., Wetselaar, A., Pees, E. & Wijffelman, C. A. *Pl. Sci. Lett.* 27, 317-325 (1982).
30. Long, S. R., Buikema, W. J. & Ausubel, R. M. *Nature* 298, 485-488 (1982).
31. Debellé, F. & Sharma, S. B. *Nucleic Acids Res.* 14, 7453-7472 (1986).

Targeting of bacterial chloramphenicol acetyltransferase to mitochondria in transgenic plants

Marc Boutry*†, Ferenc Nagy*, Carsten Poulsen*, Kazuko Aoyagi* & Nam-Hai Chua*

* Laboratory of Plant Molecular Biology, The Rockefeller University, 1230 York Avenue, New York, New York 10021-6399, USA

† Laboratoire d'Enzymologie, University of Louvain, Place Croix du Sud 1, Box 8, B-1348 Louvain-la Neuve, Belgium

Most mitochondrial proteins are encoded by nuclear genes and are synthesized as precursors containing a presequence at the N terminus. In yeast and in mammalian cells, the function of the presequence in mitochondrial targeting has been revealed by chimaeric gene studies. Fusion of a mitochondrial presequence to a foreign protein coding sequence enables the protein to be imported into mitochondria *in vitro* as well as *in vivo*¹⁻⁴. Whether plant mitochondrial presequences function in the same way has been unknown. We have previously isolated and characterized a nuclear gene (*atp2-1*) from *Nicotiana plumbaginifolia* that encodes the β -subunit of the mitochondrial ATP synthase⁵. We have constructed a chimaeric gene comprising a putative *atp2-1* presequence fused to the bacterial chloramphenicol acetyltransferase (CAT) coding sequence and introduced it into the tobacco genome. We report here that a segment of 90 amino acids of the N terminus of the β -subunit precursor is sufficient for the specific targeting of the CAT protein to mitochondria in transgenic plants. Our results demonstrate a high specificity for organelle targeting in plant cells.

Nucleotide sequence analysis revealed that the β -subunit of *N. plumbaginifolia* mitochondrial ATP synthase is synthesized as a precursor of relative molecular mass 59,000 (M_r , 59K). Because the mature β -subunit is of M_r 50K, these results together suggest a presequence of ~9K (ref. 5).

A fragment containing the putative presequence of the *atp2-1* gene was obtained by 3' *Bal31* deletions. To retain the putative

processing site, we chose a 3'-deleted fragment that encodes the amino-terminal 90 residues (M_r ~9K) of the β -subunit precursor as likely to contain the prebeta sequence as well as a few amino acids of the mature protein. After *NcoI* digestion at the first ATG⁵, the DNA fragment was fused to the CAT coding sequence by appropriate *HindIII* linkers (Fig. 1a). To express this chimaeric coding sequence (prebeta-CAT) in plant cells, it was placed under the control of the constitutive 35S promoter of the cauliflower mosaic virus (CaMV)⁶. The CAT gene without the presequence of the *atp2-1* gene and driven by the same 35S promoter⁷ was used as a negative control. Both constructs were cloned into the binary vector pMON505 (ref. 8) and introduced into tobacco cells.

Most of the transgenic plants containing either the prebeta-CAT or the control CAT construct showed CAT activity. The control CAT plant did not show any activity in organellar fractions. In contrast, CAT activity was highly enriched in the mitochondrial fraction and was barely detectable in the chloroplast fraction (Fig. 1b) of transgenic plants containing the prebeta-CAT construct.

To obtain further evidence that the prebeta-CAT product was localized in the mitochondria, we compared the CAT activity in different subcellular fractions with that of malate dehydrogenase, a marker enzyme for the mitochondrial matrix⁹. Table 1 shows that both enzymes behaved similarly during cell fractionation. About 30% of both activities in the homogenate was recovered in the crude mitochondrial fraction and nearly 50% was found in the crude supernatant. Further centrifugation of the supernatant at 100,000g for 30 min did not pellet any of the two enzyme activities (Table 1). We concluded that the soluble CAT activity could be attributed to enzymes released from mitochondria broken during homogenization.

We were unable to demonstrate directly that the prebeta-CAT product was sequestered by mitochondria because CAT was destroyed by proteolytic inactivation (not shown). However, washing mitochondria with high salt (1 M NaCl) failed to release the CAT activity (Fig. 1c) but repeated 'freeze and thaw' or sonication of the mitochondrial fraction solubilized the activity (Fig. 1c). Taken together, these results strongly suggest that the prebeta-CAT product was localized in the mitochondrial matrix.

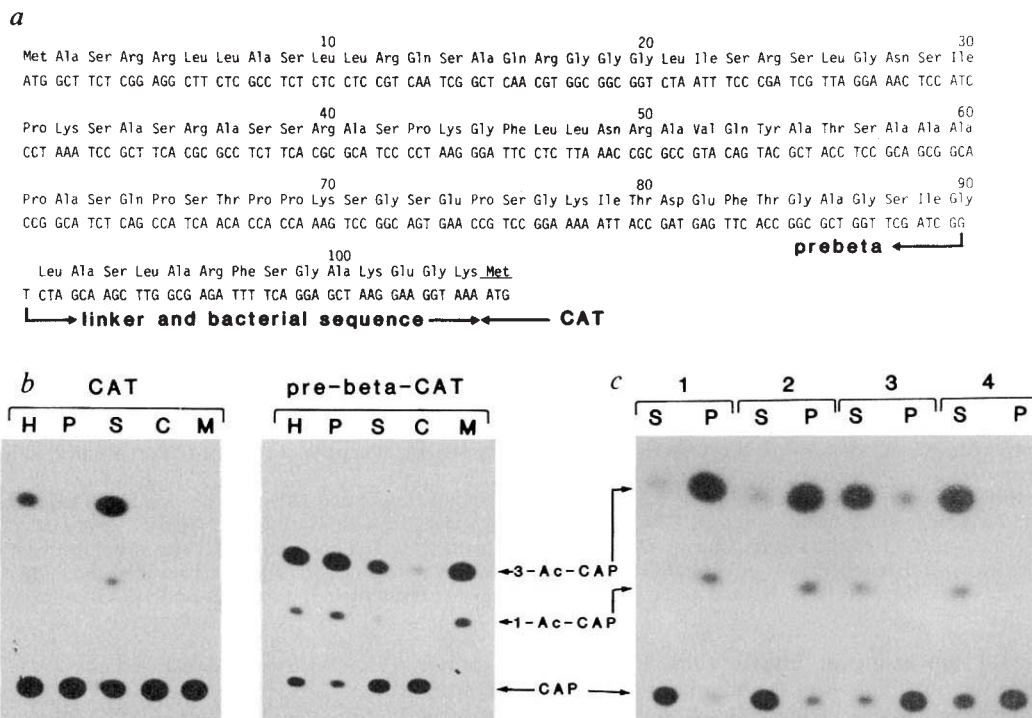
To see whether the prebeta-CAT was processed to a mature form, we used an antiserum directed against CAT for Western blot analysis. The crude cytosolic fraction of control CAT plants showed a major band of 24K (Fig. 2a), an M_r consistent with that of the bacterial enzyme¹⁰. Distribution of the prebeta-CAT products (Fig. 2b) followed that of the CAT activity (Fig. 1b). Polypeptide bands decorated by the antiserum were detected in the mitochondrial but not the chloroplast fraction. The major band (M_r = 25.5K) was slightly larger than the normal CAT protein (Fig. 2d). This apparent difference (1.5 K) probably corresponds to the linker region and to beta polypeptide residues located downstream of the processing site (see Fig. 1a). We also observed minor bands that were larger (M_r = 29K and 30K) whose intensity was slightly variable among extracts from different transgenic plants (not shown). These polypeptides were smaller than the prebeta-CAT precursor (M_r = 34.5 K) as synthesized in an *in vitro* system (Fig. 2d) and probably resulted from

Table 1 Co-fractionation of CAT activity with malate dehydrogenase activity in a prebeta-CAT transgenic plant

		Homogenate		Crude chloroplasts	Purified chloroplasts	Crude mitochondria	Purified mitochondria
CAT	Specific activity	1.00	0.75	0.55	0.25	2.8	10.9
	Proportion of homogenate (%)	100	51.7	8.3	1.0	33.3	7.4
Malate dehydrogenase	Specific activity	1.00	0.85	0.54	0.25	2.4	9.1
	Proportion of homogenate (%)	100	59.1	8.2	1.0	28.8	6.2

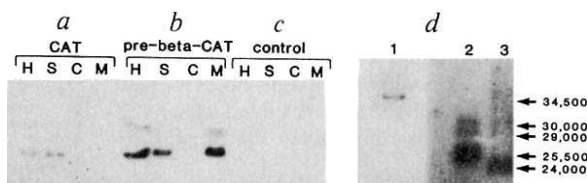
CAT and malate dehydrogenase activities were analysed in each subcellular fraction obtained from a prebeta-CAT transgenic plant as described in Fig. 1. Specific activities are expressed in arbitrary units and are given the value of 1 for the homogenate. The total activity of the fraction is given as a percentage of the total activity in the homogenate.

Fig. 1 a, The nucleotide sequence of the 5' coding region of the hybrid pre-beta-CAT gene in the expression vector pDS5. The N-terminal peptide extension of 104 amino-acid residues is composed of 90 amino acids from the beta subunit precursor and 14 amino acids from the linker region and bacterial leader sequence upstream of the CAT coding sequence¹⁵. **b**, Distribution of CAT activity in different subcellular fractions of transgenic plants with either the control 35S-CAT or the 35S-prebeta-CAT construct. CAP, chloramphenicol; 3-Ac-CAP, 3-acetyl-chloramphenicol; 1-Ac-CAP, 1-acetyl-chloramphenicol; H, homogenate; S, cytosolic supernatant; P, crude organellar pellet; M, purified mitochondria; C, purified chloroplasts. **c**, CAT activity localized in the soluble compartment of the mitochondrial fraction. Arrows, positions of labelled chloramphenicol (CA) and acetyl chloramphenicol (Ac-CA). Mitochondria were purified from a 35S-prebeta-CAT transgenic tobacco plant as in **b**. The purified fraction containing 40 µg protein in 400 µl mannitol medium was divided into four equal portions of 100 µl. The first was centrifuged for 30 min at 100,000g in a Beckman Airfuge. The supernatant was saved and the pellet (P) was resuspended in 100 µl mannitol medium. The second was adjusted to 1 M NaCl, incubated for 5 min at 4 °C and centrifuged for 30 min at 100,000g to obtain supernatant and pellet. The third was subjected to five cycles of freezing (-20 °C) and thawing. After treatment, the sample was centrifuged again as for the first two. The fourth was sonicated twice for 5 s each time in a Virsonic Disrupter and centrifuged as for the other three. From each fraction, 3 µl was assayed for CAT activity.



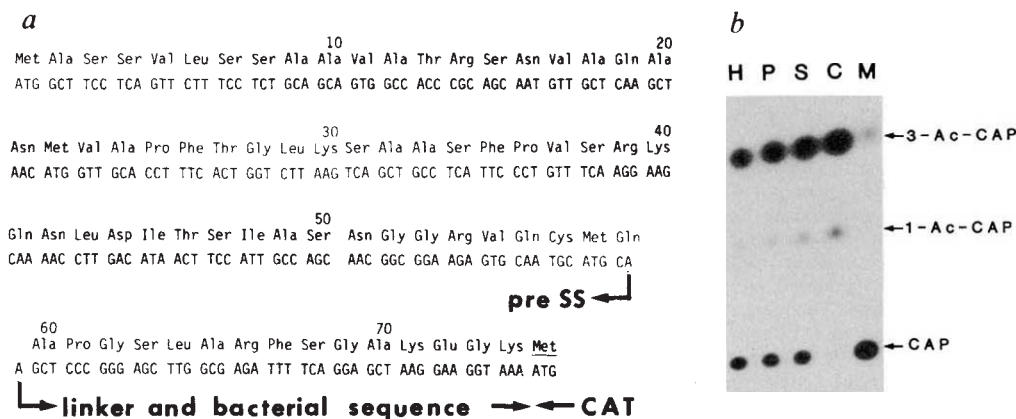
Methods. The hybrid prebeta-CAT gene was constructed as follows. A 4,803-base pair (bp) *Hind*III fragment containing the promoter region and most of the transcribed region of the *atp2-1* gene⁵ was subcloned in pEMBL8 (ref. 16). This construct was digested with *Bal*31 from the 3' end and deletion clones were sequenced after addition of *Xba*I linkers. A clone retaining 270 bp of the 5' coding region was digested with *Xba*I and at the *Nco*I site, which contains the first ATG of the *atp2-1* reading frame. The fragment was filled in with Klenow DNA polymerase and *Hind*III linkers were attached. After *Hind*III digestion the fragment was recloned into the *Hind*III site located 34 bp upstream of the bacterial CAT reading frame and 9 bp downstream of the CaMV 35S transcription initiation site⁶. In a parallel construction the prebeta fragment was introduced into the expression vector pDS5, which under the control of the coliphage T5 promoter, allowed coupled *in vitro* transcription-translation of prebeta-CAT (see Fig. 3d). For testing distribution of CAT activity, leaf material (5 g) was homogenized in a blender with 25 µl 0.33 M sucrose, 50 mM Tris HCl (pH 8.0), 0.2% bovine serum albumin (BSA), 0.1% ascorbic acid, and 0.05% β-mercaptoethanol. After filtration through four layers of Miracloth, 1 ml of the homogenate (H) was centrifuged for 10 min in an Eppendorf centrifuge to give a crude cytosolic supernatant (S) and a crude organellar pellet (P). The remaining 4.0 ml of the homogenate were centrifuged at 6,000 r.p.m. for 30 s in a Sorvall SS34 rotor. The pellet (crude chloroplast fraction) was resuspended in 2 ml of suspension medium (0.4 M mannitol, 10 mM K₂HPO₄, (pH 7.2), 0.1% BSA) layered onto a two-step Percoll gradient (4 ml 80% Percoll, 0.1% BSA and 5 ml 40% Percoll, 0.1% BSA) and centrifuged for 10 min at 10,000 r.p.m. in a Sorvall HB4 rotor. The interface which contained purified chloroplasts (C) was recovered. The supernatant from the first differential centrifugation was further centrifuged at 15,000 r.p.m. for 12 min in the Sorvall SS34 rotor. The pellet (crude mitochondrial fraction) was resuspended in 2 ml of the suspension medium, layered onto a two-step Percoll gradient (3 ml 40% Percoll, 0.1% BSA and 4.5 ml 21% Percoll, 0.1% BSA) and centrifuged for 30 min at 20,000 r.p.m. in a Beckman SW28 rotor. The interface which contained purified mitochondria (M) was recovered. Protein (1 µg) from each fraction was used for the CAT assay. CAT activity was assayed at 37 °C in 30 µl of 50 mM Tris (pH 8.0), 1 mM acetyl-CoA and 0.5 µCi ¹⁴C-chloramphenicol for 1 h. Acetylated products were identified according to Gorman *et al.*¹⁷ and quantitated as described by Suissa¹⁸.

Fig. 2 a-c, Western blot analysis of the CAT protein in transgenic tobacco: **a**, 35S-CAT control; **b**, 35S-prebeta-CAT; **c**, untransformed tobacco; H, homogenate; S, crude cytosolic supernatant; C, purified chloroplast; M, purified mitochondria. **d**, Lane 1, the prebeta-CAT protein synthesized in coupled *in vitro* transcription translation system using the plasmid pDS5 prebeta-CAT (described in Fig. 1a) as a template; lane 2, a mitochondrial fraction from a 35S-prebeta-CAT transgenic plant; lane 3, a crude cytosolic supernatant from a 35S-CAT transgenic plant. Arrows, apparent *M_r* of the visualized bands.



Methods. Subcellular fractions were obtained as in Fig. 1b. For analysis, 30 µg of each of the homogenate and the crude cytosolic supernatant or 8.0 µg of the chloroplast or the mitochondrial fractions were separated by SDS gel electrophoresis and transferred to a nitrocellulose filter. After preincubation for 30 min in 8% glycine, 1% BSA, the filters were incubated overnight at 37 °C in 150 mM NaCl, 10 mM Na₂HPO₄ (pH 7.3), 0.1% Nonidet P-40, 0.02% sodium azide, 0.3% low fat dry milk and 0.5% antiserum raised against CAT purified from bacteria. Protein blots were washed three times for 10 min each in the same medium without antiserum and then incubated with ¹²⁵I-labelled protein A as described by the supplier (Amersham). Radioactive bands were visualized by autoradiography.

Fig. 3 *a*, Nucleotide sequence of the 5' coding region of the hybrid *rbcs-8B-preSS-CAT* gene in pUC13. The amino-terminal peptide extension (74 amino-acid residues) is comprised of 59 amino acids encoded by the first exon of *rbcs-8B* (ref. 14) and 15 amino acids from the linker region and bacterial leader sequence upstream of the CAT coding sequence¹⁵. *b*, CAT activity in subcellular fractions of an *rbcs-8B-preSS-CAT* transgenic *Nicotiana plumbaginifolia*. Subcellular fractions were prepared and CAT activities measured using 1.0 µg protein as in Fig. 1*b*. Arrows, CAT and acetylated CAT. H, homogenate; S, cytosolic supernatant; P, crude organellar pellet; C, purified chloroplast fraction; M, purified mitochondrial fraction.



Methods. A *Bam*HI-*Sph*I fragment containing the 1,038-bp promoter region and 250 bp of the transcribed sequence from the *N. plumbaginifolia* *rbcs-8B* gene¹⁴ was subcloned into pUC19. The plasmid was digested with *Hind*III and *Hind*III-*Sma*I adapters were attached by linker ligation. A pUC13 plasmid containing a CaMV-35S-CAT construct¹⁹ with a unique *Hind*III site 34 bp upstream of the first ATG of the CAT gene was also digested with *Hind*III and *Hind*III-*Sma*I adapters were attached. The *Sma*I-adapted *rbcs-8B* sequence was then introduced into the *Sma*I-adapted CAT plasmid as an *Eco*RI-*Sma*I fragment resulting in the *rbcs-8B-preSS-CAT* chimaeric gene.

abnormal processing at cryptic sites. However, we cannot exclude the possibility that the minor bands were derived from correct precursor processing, but due to their instability were cleaved further to yield the *M_r* 25.5K band.

These results strongly suggest a strict specificity of the β -subunit presequence for mitochondrial targeting. The specificity of organelle presequences, however, has been challenged recently in a set of experiments demonstrating that the *Chlamydomonas* *rbcs* transit sequence can direct transport of a foreign protein into yeast mitochondria, although the process is less efficient than that mediated by a mitochondrial presequence¹¹. The pea *rbcs* transit sequence was used previously to target neomycin phosphotransferase into chloroplasts^{12,13}, but the specificity of organelle targeting was not analysed. To address this question we designed a hybrid gene in which the CAT coding sequence was fused to the chloroplast transit sequence of *rbcs8B* from *N. plumbaginifolia*¹⁴ (Fig. 3*a*). The *rbcs-preSS-CAT* hybrid gene was introduced into *N. plumbaginifolia* and subcellular fractions of transgenic plants were analysed. Figure 3*b* shows that CAT activity was associated with the purified chloroplast fraction. Enrichment from the homogenate was no more than twofold but this was not unexpected given the abundance of chloroplast proteins in leaf cells. On the other hand the low CAT activity in the mitochondrial fraction from these plants indicated the specificity of the chloroplast transit sequence for the corresponding organelle. Because the same reporter protein was used for targeting to mitochondria and to chloroplast, our results demonstrate unambiguously that the presequence determines the specificity in organelle targeting and that the structure of the 'passenger' protein we studied is not involved.

In this paper we report the first analysis of a plant presequence involved in mitochondrial targeting. With this technique, it is now possible to introduce novel traits that would be expressed specifically in plant mitochondria so that unresolved aspects of mitochondrial biogenesis and molecular aspects of male sterility in higher plants can be studied.

We thank I. Roberson, M.-Y. Hsu, M. Wong and A.-M. Faber-Hubermont for excellent technical help, Dr B. Dobberstein for pDS5, J. O'Donnell for graphics and W. Roine for typing. M.B. thanks Drs A. Goffeau and M. Briquet (University of Louvain) for their support and interest. M.B. is a Research Associate at the Federation Nationale de Recherche Scientifique (Belgium). Work done at The Rockefeller University was supported by grants from Monsanto, NIH, and a Winston Foundation Fellowship to M.B.

Received 12 March; accepted 5 June 1987.

- Emr, S. et al. *J. Cell Biol.* **102**, 523-533 (1986).
- Horwich, A., Kalousek, F., Mellman, I. & Rosenberg, I. *EMBO J.* **4**, 1129-1135 (1985).
- Hurt, E., Muller, U. & Schatz, G. *EMBO J.* **4**, 3509-3518 (1985).
- Van Loon, A. P., Brandli, A. W. & Schatz, G. *Cell* **44**, 801-812 (1986).
- Boutry, M. & Chua, N.-H. *EMBO J.* **4**, 2159-2165 (1985).
- Odell, J. T., Nagy, F. & Chua, N.-H. *Nature* **313**, 810-812 (1985).
- Morelli, G., Nagy, F., Fraley, R. T., Rogers, S. G. & Chua, N.-H. *Nature* **315**, 200-204 (1985).
- Horsch, R. B. & Klee, H. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4428-4432 (1986).
- Ochoa, S. *Meth. Enzym.* **1**, 735-739 (1955).
- Shaw, W. V. *CRC crit. Rev. Biochem.* **14**, 1-46 (1982).
- Hurt, E., Soltanifar, N., Goldschmidt-Clermont, M., Rochaix, J.-D. & Schatz, G. *EMBO J.* **4**, 1343-1350 (1986).
- Van den Broeck, G. et al. *Nature* **313**, 358-363 (1985).
- Schreier, P., Seftor, E. A., Schell, J. & Bohnert, H. J. *EMBO J.* **5**, 1335-1342 (1985).
- Poulsen, C., Fluhr, R., Kauffman, J. M., Boutry, M. & Chua, N.-H. *Molec. gen. Genet.* **205**, 193-200 (1986).
- Stueber, D., Ibrahim, I., Cutler, D., Dobberstein, B. & Bujart, H. *EMBO J.* **3**, 3143-3148 (1984).
- Dente, L., Cesareni, G. & Cortese, R. *Nucleic Acids Res.* **11**, 1645-1655 (1983).
- Gorman, C. M., Moffat, L. F. & Howard, B. H. *Molec. cell. Biol.* **2**, 1044-1051 (1982).
- Suissa, M. *Analyt. Biochem.* **133**, 511-514 (1983).
- Hunt, A., Chu, N., Odell, J., Nagy, F. & Chua, N.-H. *Pl. Molec. Biol.* **8**, 23-35 (1987).

Altered adhesive interactions with marrow stroma of haematopoietic progenitor cells in chronic myeloid leukaemia

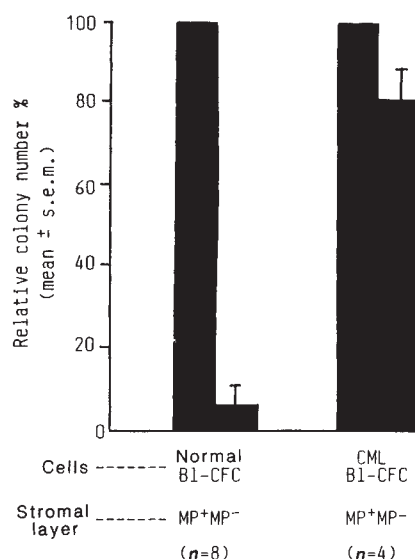
M. Y. Gordon, C. R. Dowding*, G. P. Riley, J. M. Goldman* & M. F. Greaves

The Leukaemia Research Fund Centre, at the Institute of Cancer Research, Chester Beatty Laboratories, Fulham Road, London SW3 6JB, UK

* The MRC Leukaemia Unit, Royal Postgraduate Medical School, Du Cane Road, London W12 0HS, UK

Normal haematopoietic cell regulation involves interaction between marrow stromal cells and haematopoietic progenitor cells¹ which may be facilitated by specific recognition and adhesion²⁻⁴. Some leukaemogenic events might produce a selective growth advantage by altering this regulatory network, possibly by diminishing the capacities of cells to adhere to stromal elements. Using an *in vitro* culture system which allows investigation of adhesion to stromal layers and subsequent colony formation by blast colony-forming cells (B1-CFC) in normal marrow²⁻⁴ and Ph⁺ chronic myeloid leukaemic (CML) blood⁵, we compared the

Fig. 1 Comparison of colony formation on MP^- stromal layers with colony formation on MP^+ stromal layers. The stroma layers were produced by plating 5×10^5 lymphoprep-separated (Nyegaard) normal marrow mononuclear cells in 1 ml α -medium (GIBCO) with 15% fetal calf serum (FCS; GIBCO) and with (MP^+) or without (MP^-) 2×10^{-6} M methylprednisolone (Upjohn) in 35-mm Petri dishes. Stromal cultures were fed weekly by complete replacement of medium and supplements until confluent. For the B1-CFC assay, 5×10^5 normal marrow or CML blood mononuclear cells, pre-depleted by plastic-adhesion, were added to each stromal layer for 2 h co-incubation. The stromal layers were then washed thoroughly and overlaid with 1 ml 0.3% agar in α -medium + 15% FCS. Colonies of more than 20 cells were scored and identified by staining with May Grunwald-Giemsa after 5 days incubation at 37 °C in humidified 5% CO_2 in air. The results are expressed as the numbers of colonies forming on MP^- stromal layers relative to the numbers on MP^+ stromal layers. The vertical bars represent the standard errors of the interexperimental means; n is the number of replicate experiments.



adhesive properties of normal and malignant progenitor cells. We present evidence that altered adhesive interactions between primitive progenitor cells and marrow stromal cells occur in CML.

In CML, clonal dominance of transformed pluripotent stem cells⁶ is associated with enhanced *c-abl* protein kinase activity resulting from the Philadelphia chromosome translocation⁷. Expansion of the leukaemic clone and premature release of Ph^+ haematopoietic progenitor cells from the marrow into the blood leads to haematopoiesis in extramedullary sites. Paradoxically, in long-term cultures of marrow from CML patients the Ph^- progenitors (which may be normal) have a selective advantage⁸. This clonal advantage, premature release, extramedullary expansion and negative selection of leukaemic cells *in vitro* might all be explained by altered adhesive interactions with stromal elements.

Putative stem cells in normal human marrow can be assayed by a blast colony-forming assay which requires recognition and adhesion to a pre-formed stromal layer²⁻⁴. B1-CFC are not normally released into the bloodstream² but numerous B1-CFC circulate in CML⁵.

We compared B1-CFC in CML blood with B1-CFC in normal marrow because it is the leukaemic counterparts of these normal progenitors which circulate in CML⁹. Normal B1-CFC adhere to and form colonies on stromal cells cultured with 2×10^{-6} M methylprednisolone ($=MP^+$ stroma) but not on MP^- (steroid-untreated) stroma (Fig. 1; ref. 4). Presumably this difference reflects specific requirements for haematopoiesis *in vivo*. In marked contrast, B1-CFC from CML blood can circumvent this requirement and form colonies equally well on MP^+ and MP^- stroma *in vitro* (Fig. 1).

The adhesion of B1-CFC to cultured stroma can be quantified by measuring loss of cells from a marrow suspension during co-incubation ('panning') with a pre-formed stromal layer⁴. Over 90% of normal B1-CFC are removed by panning on an MP^+ stromal layer; none are removed on MP^- stromal layers (Fig. 2a). In contrast, normal granulocyte-macrophage colony-forming cells (GM-CFC), which are the progeny of B1-CFC³, are not adhesive (Fig. 2b). CML B1-CFC formed colonies on both primary (panning) and secondary (assay) stromal layers resulting in a recovery of about 40% (Fig. 2c) and there was no difference in the capacity of MP^+ or MP^- stromal layers to remove B1-CFC from CML blood. The adhesion properties of CML B1-CFC are, therefore, intermediate between those of normal B1-CFC and those of normal GM-CFC. But circulating B1-CFC and GM-CFC in CML differ markedly in number^{10,11} and also in their levels, with respect to normal, of HLA-DR antigen expression (C.R.D., unpublished observations).

Because the increased recovery in CML cannot be explained

by saturation of stromal binding sites by excess progenitor cells³, the overall efficiency/avidity of B1-CFC adhesion may be reduced in CML or there may be two populations of progenitor cells (that is, binders and non-binders). To distinguish between these two possibilities we increased the co-incubation time of CML cells and MP^+ stromal layers to 6 h. This did not alter the results (Table 1). Cell-cycle related expression of adhesion sites was not responsible for the apparent heterogeneity of CML cells (see ref. 8) because hydroxyurea-treated and untreated CML

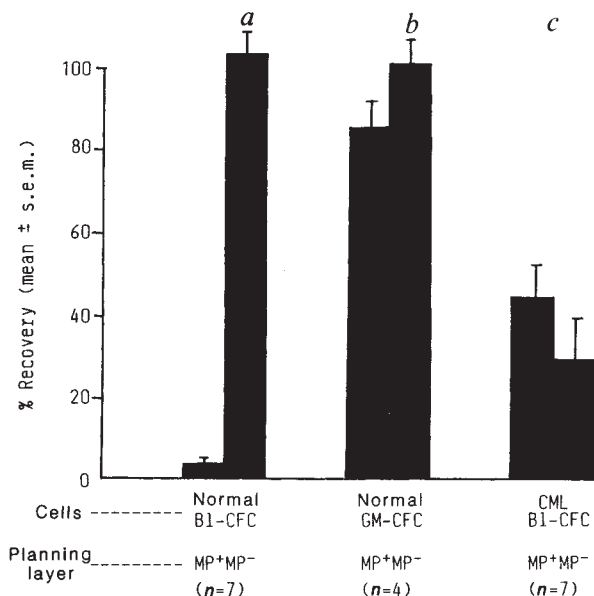


Fig. 2 Recovery of B1-CFC from normal marrow or CML blood and of normal marrow GM-CFC after panning on MP^+ or MP^- stromal layers. The stromal layers were prepared as described in the legend to Fig. 1 and cells were co-incubated with the stromal layers for 2 h. Thereafter, the cells remaining in suspension were transferred to further MP^+ stromal layers (B1-CFC) or to 0.3% agar cultures containing α -medium, 15% FCS and 15% 5637 conditioned medium as a source of colony-stimulating factor¹⁵ for bioassay of GM-CFC. The results of the B1-CFC assays are expressed as the proportion of total progenitor cells (sum of colony numbers on primary and secondary MP^+ stromal layers) producing colonies on secondary plates. The results of the GM-CFC assays are relative to the numbers of colonies grown from cell suspensions incubated in parallel in otherwise empty dishes. The vertical bars represent the standard errors of the interexperimental means; n is the number of replicate experiments.

Table 1 Adhesion by CML-B1-CFC in relation to duration of incubation, exposure to hydroxyurea and number of panning steps

Colonies per plate/Percentage recovered after panning				
Time (h)				
0.5	(a) 153 ± 19/49 ± 7%	(c) 82 ± 2/41 ± 2%	(d) ND	
2.0	224 ± 26/40 ± 5%	91 ± 4/33 ± 2%	115 ± 4/49	
			± 4	
4.0	257 ± 17/35 ± 2%	99 ± 3/31 ± 1%	111 ± 5/43	
			± 2	
6.0	245 ± 17/38 ± 2%	97 ± 4/32 ± 2%	84 ± 2/48	
			± 1	
Hydroxyurea*				
Control	(b) 99 ± 2/45 ± 2%	91 ± 4/33 ± 2%	115 ± 4/49	
			± 4	
Treated	84 ± 2/42 ± 3%	93 ± 3/31 ± 1%	95 ± 4/49	
			± 3	
Triple panning†				
Primary stroma	99 ± 2/—	91 ± 4/—	115 ± 4/—	
Secondary stroma	82 ± 3/45 ± 2%	45 ± 2/33 ± 2%	110 ± 7/49 ±	
			4	
Tertiary stroma	39 ± 5/32 ± 4%	28 ± 2/39 ± 4%	61 ± 3/36 ±	
			2	

a, b, c and d; four different CML blood samples used for these experiments. Colonies per plate given as mean ± standard error of intraexperimental mean, and % recovered after panning is ± the standard error of the ratio¹⁶.

* Cells were depleted by adhesion in plastic flasks, then were treated with 2.5×10^{-3} M hydroxyurea for 1 h at 37 °C and washed three times before plating. Exposure of cells in sample d to increased concentrations (5 and 10×10^{-3} M) of hydroxyurea did not reduce the survival of the colony-forming cells. Parallel cultures of day 7 and day 14 GM-CFC from the same flask gave survival values of 70% and 63% respectively.

† Each panning interval was of 2 h duration.

progenitors adhered equally well (Table 1). Finally, a further panning step showed that the proportion of B1-CFC transferred from secondary to tertiary plates did not differ from that transferred from primary to secondary plates (Table 1). Hence, the results were not greatly influenced by altered efficiencies of adhesion at the different stages. This result suggests that the CML blast progenitor cells have a reduced capacity to adhere to MP⁺ stroma. The apparent heterogeneity is most probably a consequence of the assay which requires an arbitrary threshold for binding.

Our results reveal two potentially important differences between B1-CFC in normal marrow and those in CML blood. First, the CML cells have an abnormal capacity to adhere to MP⁺ stroma, indicating that the cells are less discriminating than normal. This could provide a growth advantage for leukaemic cells *in vivo* by increasing the extramedullary space (equivalent to MP⁺ stroma *in vitro*) available for leukaemic progenitor cell expansion. Second, the CML cells are deficient in their adhesion to the MP⁺ stroma. As we use normal marrow stromal cells in these experiments, we conclude that there is an intrinsic abnormality in the capacity of primitive haematopoietic progenitors in CML to interact with stromal elements.

If adhesive interactions with marrow stromal cells retain normal stem cells in the marrow, the inefficient adhesion to stroma exhibited by CML cells *in vitro* could account for their premature release into the blood. It could also account for the paradox that long-term marrow cultures, in which normal cells proliferate within the stromal layer, offer a growth advantage to Ph⁻ cells in CML⁸. This observation, which has implications for

autologous marrow transplantation¹², can now be explained by an increased release of proliferating Ph⁺ cells¹³ and, hence, greater loss of leukaemic cells when the cultures are fed.

The structure on normal marrow B1-CFC which recognizes cultured stroma has not yet been identified. But it is not sialation of this structure that causes abnormal adhesion by CML cells (see Baker *et al.*¹⁴) because neuraminidase-treatment of normal or CML cells did not alter their adhesion to MP⁺ or MP⁻ stroma (four experiments—data not shown). Alternatively, as normal marrow B1-CFC may recognize heparan sulphate produced by marrow stromal cells (M.Y.G., unpublished observations) the data imply that the capacity of B1-CFC in CML to recognize this structure is diminished.

We cannot yet determine whether altered adhesive interactions with stroma in CML are causal or a secondary consequence. Although more than one genetic mechanism may contribute to the development of CML, the enhanced *c-abl* kinase activity⁷ is presumably crucial because it is invariably associated with CML. The substrate for the altered *c-abl* kinase is unknown but it may be a cell-surface or cytoskeletal protein involved in cell-cell adhesion. It is of some interest in this respect that phosphotyrosine-containing proteins may be concentrated in focal adhesions in normal cells¹⁷ and that viral pp60^{src} may be localized at sites of fibronectin degradation¹⁸.

We thank the bone marrow transplantation teams of the Royal Marsden and Hammersmith Hospitals for supplying the bone marrow samples used in these experiments and Miss Geraldine Parkins for expert preparation of this manuscript. This work was supported by the Leukaemia Research Fund.

Received 16 March; accepted 5 June 1987.

- Dexter, T. M. *Clin. Haemat.* **8**, 453–468 (1979).
- Gordon, M. Y., Hibbin, J. A., Kearney, L. U., Gordon-Smith, E. C. & Goldman, J. M. *Br. J. Haemat.* **60**, 129–136 (1985).
- Gordon, M. Y., Dowding, C. R., Riley, G. P. & Greaves, M. F. *J. Cell Physiol.* **130**, 150–156 (1986).
- Gordon, M. Y., Hibbin, J. A., Dowding, C., Gordon-Smith, E. C. & Goldman, J. M. *Expl. Hematol.* **13**, 937–940 (1985).
- Dowding, C. R., Gordon, M. Y., Goldman, J. M. *Int. J. Cell Cloning* **4**, 331–340 (1986).
- Falkow, P. J. *J. Cell Physiol.* (Suppl. 1) 37–43 (1982).
- Konopka, D. B., Wanatabe, S. M., Singer, J. W., Collins, S. J. & Witte, O. N. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1810–1814 (1985).

- Coulombel, L., Kalousek, D. K., Eaves, C. J., Gupta, C. M. & Eaves, A. C. *New Engl. J. Med.* **308**, 1493–1498 (1983).
- Robak, T., Nolasco, I., Hibbin, J. & Goldman, J. M. *Leukaemia Res.* **9**, 1023–1029 (1985).
- Goldman, J. M., Shiota, F., Th'ng, K. H. & Orchard, K. H. *Br. J. Haemat.* **46**, 7–13 (1980).
- Dowding, C. R., Gordon, M. Y. & Goldman, J. M. *Int. J. Cell Cloning* **4**, 331–340 (1986).
- Chang, J. *et al. Lancet* **i**, 294–295 (1986).
- Eaves, A. C., Cashman, J. D., Gaboury, L. A., Kalousek, D. K. & Eaves, C. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5306–5310 (1986).
- Baker, M. A., Taub, R. N., Kanani, A., Brockhausen, I. & Hindenberg, A. *Blood* **44**, 1068–1071 (1985).
- Myers, C. D., Katz, F. E., Joshi, G. & Millar, J. L. *Blood* **64**, 152–155 (1984).
- Blackett, N. M. *Cell Tissue Kinetics* **7**, 407–412 (1974).
- Maher, P. A., Pasquala, E. B., Wang, J. Y. J. & Singer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6576–6580 (1985).
- Chen, W.-T., Chen, J.-M., Parsons, S. J. & Parsons, J. T. *Nature* **316**, 156–158 (1985).

HIV-specific cytotoxic T lymphocytes in seropositive individuals

Bruce D. Walker*, Sekhar Chakrabarti†, Bernard Moss†, Timothy J. Paradis*, Theresa Flynn*, Amy G. Durno*, Richard S. Blumberg*, Joan C. Kaplan*, Martin S. Hirsch* & Robert T. Schooley*

* Infectious Disease Unit, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114, USA

† Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

Virus-specific cytotoxic T lymphocytes (CTL) which kill virus-infected cells are thought to be a major host defence against viral infections¹⁻¹¹. Here we report the existence of human immunodeficiency virus (HIV)-specific CTL in persons infected with this virus, the aetiological agent of AIDS (acquired immunodeficiency syndrome). Recombinant HIV-vaccinia viruses were used to express HIV antigens in B-cell lines established from subjects seropositive for HIV and seronegative controls. Circulating lymphocytes capable of killing HIV *env*-expressing autologous B cells were detected in eight of eight seropositive subjects; in addition, at least three seropositive subjects demonstrated *gag*-specific cytotoxic responses. No HIV-specific cytotoxicity was observed in seronegative subjects. Selective inhibition of the *env*-specific cytotoxicity by a CD3-specific monoclonal antibody indicates that the effectors are T cells. This demonstration of a cytotoxic T-cell immune response to HIV in infected individuals should prove useful in investigating the immunopathogenesis of HIV infection further and in evaluating AIDS vaccine strategies.

Recombinant vaccinia viruses expressing foreign viral gene products have been used to characterize virus-specific cellular immune responses in a number of viral systems¹²⁻¹⁸. Recently, the construction of a recombinant vaccinia virus vector expressing the HIV *env* gene was described^{19,20}. Infection of susceptible cells with this recombinant results in normal synthesis, glycosylation, processing and membrane transport of the *env* polypeptide, which can be recognized by serum antibodies from patients with AIDS¹⁹. The gene product has been shown to be functional by demonstration of syncytia formation with cells expressing the CD4 cell-surface epitope²¹. These features suggested that recombinant vaccinia vectors might be useful in making target cells which express HIV antigens for the study of cell-mediated immunity to the AIDS virus. As vaccinia-infected B cells have been shown to be sensitive target cells¹⁶, we developed B lymphoblastoid lines from HIV seropositive subjects and seronegative controls by infecting peripheral blood mononuclear cells (PBMC) with Epstein-Barr virus (EBV), as previously described²⁴. These cell lines were then infected with HIV-vaccinia recombinant viruses and used as target cells in a chromium-release assay.

Three recombinant vaccinia vectors were used in these experiments. VAC/*env* (referred to elsewhere as VSC-25) expresses the HIV *env* glycoprotein in the absence of other HIV structural or regulatory proteins¹⁹. The VAC/*gag* recombinant (VSC-40), constructed in similar fashion to VAC/*env*, contains the entire *gag-pol* region of the HIV genome, extending from an *Sst*I site at nucleotide 226 to an *Nde*I site at base 4,706 (ref. 22). This vector has been shown by radioimmunoprecipitation to express the p55 *gag* structural protein, but further processing to p24 has not been seen, nor has reverse transcriptase activity been detected. VAC/*lac* (VSC-8), a recombinant vaccinia vector containing the bacterial *lacZ* gene, was used as a control.

Initially, we evaluated the effects of infection of immortalized B cells with the HIV-vaccinia recombinant vectors, and deter-

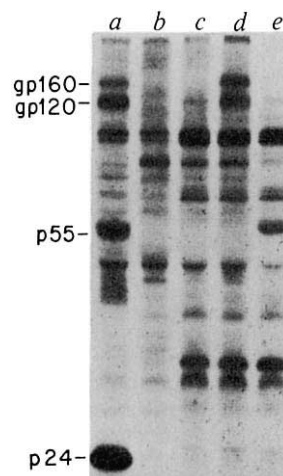


Fig. 1 Demonstration by radioimmunoprecipitation of HIV gene products in B cells infected with HIV-vaccinia recombinant vectors. *a*, HTLV-IIIIB-infected H9 cells; *b*, mock-infected B lymphoblast control; *c*, B lymphoblasts infected with VAC/*lac*; *d*, infected with VAC/*env*; and *e*, infected with VAC/*gag*. The positions of the gp160 and gp120 envelope proteins, as well as the p55 and p24 *gag* proteins, are shown. Although p55 is observed in lane *e*, there is no evidence of further processing to p24. A band migrating midway between p55 and p24 in the vaccinia-infected cells was also precipitated by HIV-seronegative serum (not shown).

Methods. Isolation of peripheral blood lymphocytes, immortalization of B cells and maintenance of cultures were performed as described previously²⁴, except that in most cases cells were maintained in RPMI 1640 media supplemented with 20% rather than 10% fetal calf serum (FCS). All established B-cell lines were free of reverse transcriptase activity, assayed as previously described³⁰. Vaccinia infection of the EBV-immortalized B cells was performed using a modification of a previously described procedure¹⁶. Log phase lymphoblasts (10×10^6) were suspended in 300 μ l of media containing 10 plaque-forming units (p.f.u.) per cell of the appropriate vaccinia recombinant and incubated for 1.5 h at 37 °C. Mock infection was carried out in similar fashion except that no vaccinia virus was added. These acutely infected cells were then washed and either resuspended at 2×10^6 cells ml^{-1} in fresh medium for a 16-h overnight incubation, then chromium labelled (see Fig. 2 legend), or resuspended at 1×10^6 per ml in 2 ml cysteine- and methionine-free medium, to which radiolabelled cysteine and methionine had been added. Following a 16-h incubation, cells were immunoprecipitated as described³¹ using serum from an HIV-seropositive subject. An autoradiograph of a polyacrylamide gel is shown.

mined that these viruses induced synthesis of the desired immunologically reactive HIV proteins in B cells (Fig. 1). In addition, the biological activity of the *env* protein was demonstrated by syncytium formation when the VAC/*env*-infected B cells were cocultivated with C8166 cells²³, a strongly CD4⁺ cell line (see legend, Fig. 2).

We then used these cells as targets in an autologous chromium-release assay. In this assay, lysis of labelled target cells results in release of chromium into the medium, providing an indirect measure of cytotoxicity. B-cell lines were established from eight HIV-seropositive homosexual males and five seronegative controls and infected with the recombinant vaccinia vectors. Freshly isolated autologous PBMCs were used as effector cells in the assay. The results for each subject, grouped by clinical and serological status, are presented in Fig. 2 at a representative effector: target (E:T) ratio of 50:1. No consistent HIV-specific response was observed in seronegative individuals (Fig. 2a). However, in each of the seropositive subjects (Fig. 2b, c) the envelope-specific cytotoxic response was two to fourteen times that observed for the VAC/*lac* infected target cells. In addition, the specific killing of the VAC/*gag* infected targets was consistently higher than the vaccinia control in seropositive subjects.

Fig. 2 Specific cytotoxicity for each subject at a fixed effector: target ratio (50:1). Subjects are grouped by serological status and disease category. *a*, Seronegative control subjects; *b*, healthy HIV-seropositive subjects; *c*, ARC/AIDS subjects. ■, EBV targets; □, VAC/*lac*-infected targets; ▨, VAC/*gag*-infected targets; ▩, VAC/*env*-infected targets. Seropositive subjects were all homosexual men and consisted of four healthy seropositive subjects, three subjects with AIDS-related complex (ARC), and one subject with AIDS. Four heterosexual seronegative laboratory workers (3 male, 1 female) and one seronegative homosexual male served as controls. Cells from three healthy seropositive subjects were assayed more than once (vertical lines represent standard error of the mean for these subjects). Aggregate results represent 19 assays in 13 subjects.

Methods. Sixteen hours after infection or mock infection, 1×10^6 cells were pelleted and resuspended in 100 μ l of normal saline containing 100 μ Ci of $\text{Na}_2^{51}\text{CrO}_4$ (New England Nuclear) for 45 min at 37 °C, washed extensively, and used as targets in a 6-h chromium-release assay, performed in triplicate as described²⁴. Percent viabilities (determined by trypan blue exclusion) of the target cells averaged for all subjects at the start of the assay were: mock-infected EBV target, 92%; VAC/*lac* infected target, 81%; VAC/*env* infected target, 81%; VAC/*gag* infected target, 82%. There were no differences in percent viability between seropositive and seronegative groups. At the time of each assay, *env* expression was verified by co-cultivating 5×10^5 VAC/*env*-infected lymphoblasts with 1.5×10^6 C8166 cells. After a 4-h incubation at 37 °C, syncytium formation was always present. Syncytia were never seen in parallel control co-cultivations using mock, VAC/*lac*- and VAC/*gag*-infected lymphoblasts. In addition, no syncytia were observed during co-cultivation of effector cells with VAC/*env*-infected target cells in the chromium release assay, although minimal cytopathic effects were observed with all vaccinia-infected target cells. Percent specific ^{51}Cr release (percent specific cytotoxicity) was determined by the formula: $100 \times [(\text{release in assay} - \text{spontaneous release}) / (\text{maximum release} - \text{spontaneous release})]$. Maximum release was determined by water lysis of labelled target cells. Average spontaneous release values for the different targets, measured by incubating target cells in the absence of effector cells, were: mock-infected B lymphoblast control, 18%; VAC/*lac*-infected control target, 19%; VAC/*env*-infected target, 18%; VAC/*gag*-infected target, 24%. There were no differences in spontaneous release between seronegative and seropositive groups. Assays were excluded from the final tabulation if the viability of the EBV line before vaccinia infection was <75% (2 assays), if the average viability following ^{51}Cr labelling was <70% (1 assay) or if the average spontaneous release was >30% (2 assays). In every case except one, repeating the assay provided data included in the final tabulation. The one subject assayed but not included in this paper's results was an AIDS patient receiving azidothymidine (AZT) therapy in whom average viability after ^{51}Cr labelling was <70%. Profound leukopenia has precluded repeating the experiment in this subject.

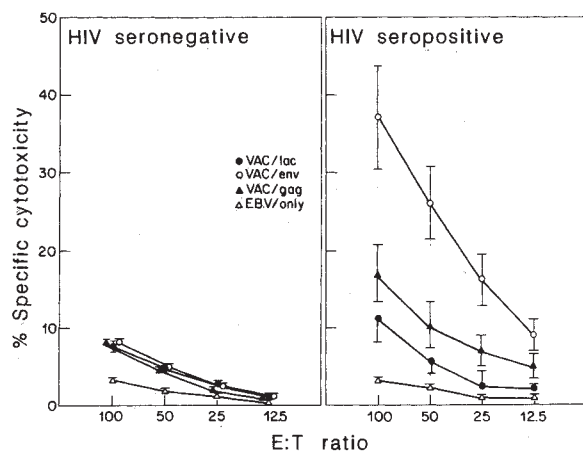
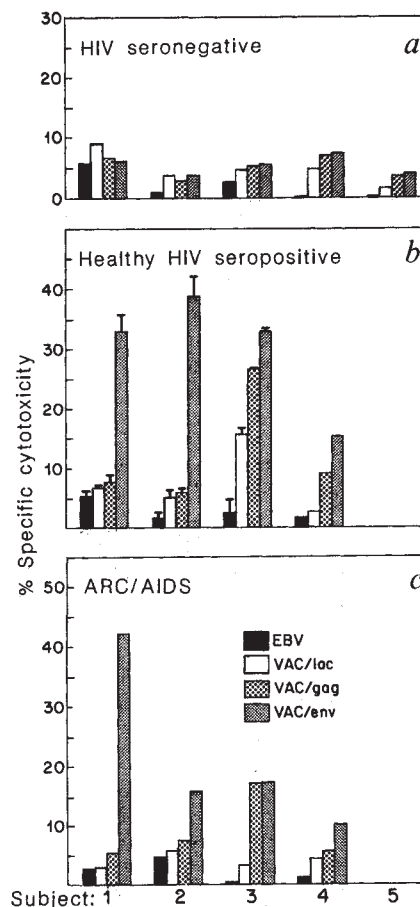
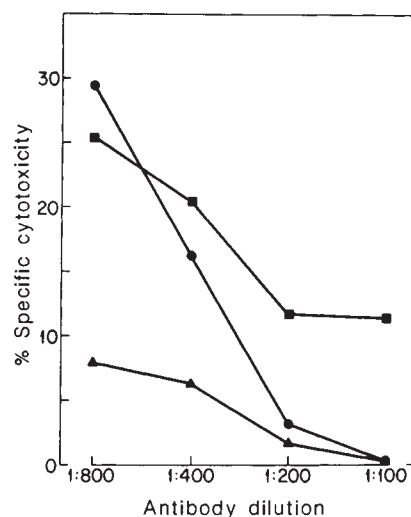


Fig. 3 Percent specific cytotoxicity averaged for seronegative and seropositive subjects. Effector: target (E:T) ratios are as noted. Results reflect the mean $\pm$ standard error (vertical bars) for all subjects assayed.

Although the *gag*-specific response was only minimally elevated in five seropositive subjects, in three subjects (Fig. 2*b* subjects 3 and 4, Fig. 2*c* subject 3) the response ranged from 1.7 to 5.5 times that of the paired VAC/*lac* controls. In the three subjects assayed more than once, the HIV-specific cytotoxicity was highly reproducible, as reflected in the standard error bars in Fig. 2*b*. Background cytotoxic response to the EBV target was low in all groups, consistent with previous observations that a period of *in vitro* sensitization is necessary to detect EBV-specific killing²⁴. The background cytotoxic response to the VAC/*lac*-infected control was low in all seronegative subjects, including two laboratory workers who had recently received a vaccinia immunization. This is consistent with previous observations that circulating vaccinia-specific cytotoxic cells disappear rapidly after vaccination²⁵. Similarly, low background cytotoxicity of the VAC/*lac*-infected control was found in all but one of the seropositive subjects (Fig. 2*b*, subject 3). The reason for augmented vaccinia-specific cytotoxicity in this healthy seropositive subject is not known.

Results were also analysed in aggregate for the eight seropositive and five seronegative subjects (Fig. 3). An *env*-specific cytotoxic response was observed in seropositive subjects at all four E:T ratios when compared with their VAC/*lac* infected controls ($P < 0.005$, Student's paired *t*-test). In addition, there was a statistically significant increase in *gag*-specific killing at

Fig. 4 Inhibition of HIV *env*-specific cytotoxicity by a monoclonal anti-CD3 antibody in three subjects. Specific cytotoxicity in the absence of added antibody was 14.0% (▲), 34.5% (●) and 38.1% (■). Antibody was incubated with effector cells in 100 μ l media at 37 °C for 30 min and 1×10^4 labelled target cells in 100 μ l medium were added, resulting in the specified antibody dilutions in 200 μ l final volume. Effector: target ratio was 50:1 in each case. Values represent average percentage specific 51 Cr release in a 6-h assay performed in triplicate. The antibody, 12F6, is a murine ascitic fluid monoclonal antibody of the same isotype and specificity as OKT₃ (ref. 32).



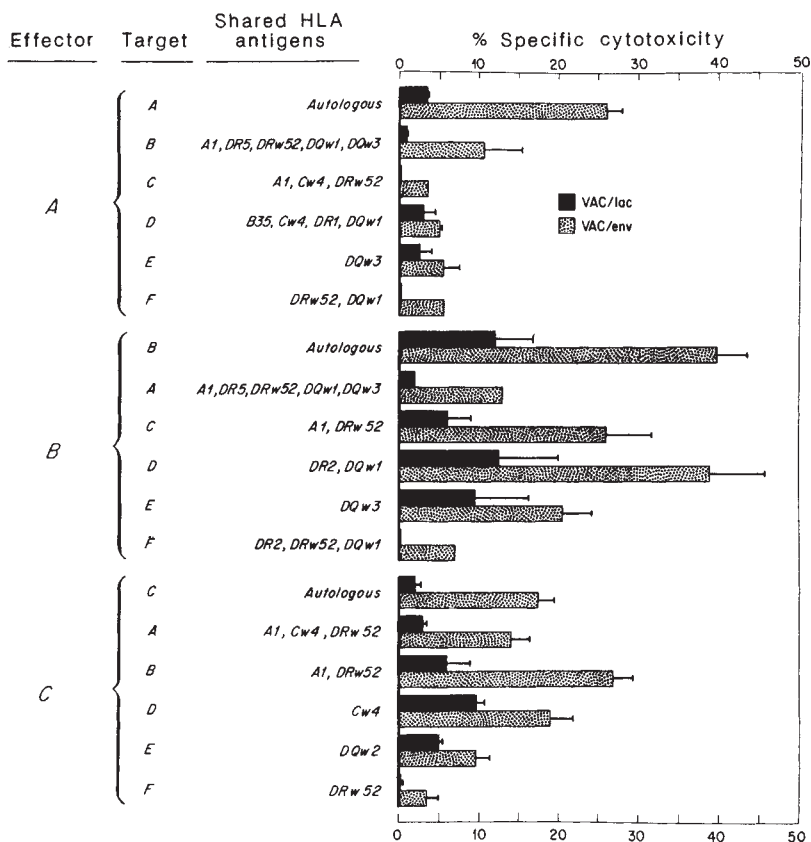
E:T ratios of 100:1, 50:1, and 25:1, compared with their VAC/*lac*-infected controls ($P < 0.05$, Student's paired t -test). Similar statistical evaluation revealed no evidence of HIV-specific cytotoxicity in seronegative individuals.

We next undertook to identify the specific cell type responsible for the observed cytotoxicity. Addition of a CD3-specific monoclonal antibody inhibited the envelope-specific cytotoxic activity in the three seropositive subjects tested, confirming that the response is T-cell mediated (Fig. 4).

In a separate set of experiments, the ability of effector cells from seropositive subjects to kill HLA-non-identical target cells expressing the HIV envelope gene was investigated (Fig. 5). In each instance in which heterologous envelope-specific killing was observed, at least one serologically-defined HLA antigen was shared between effectors and targets. In addition, effector cells from two of the three subjects tested (Fig. 5, subjects A

and B) caused preferential killing of the autologous envelope-expressing target, indicating that the major histocompatibility complex (MHC) may be important in restricting the cytotoxic response. Using effector cells from a third seropositive subject (Fig. 5, subject C), killing of three heterologous envelope-expressing targets was comparable to or greater than that of the autologous target. However, effectors from this subject also caused increased killing of the heterologous VAC/*lac*-infected control targets, indicating that this higher background cytotoxicity might account for some of the observed heterologous envelope-specific cytotoxicity. There was no killing of heterologous HIV envelope-expressing target cells using effector cells from seronegative subjects. Further delineation of the role of MHC in the cytotoxic response, as well as definition of alleles permissive for CTL lysis, will require additional investigation, ideally using cloned effector T cells.

Fig. 5 HIV-specific cytotoxicity of autologous and heterologous B-cell targets. Cytotoxicity assays were performed using effector cells from three seropositive subjects (A, B and C) with autologous and heterologous target cells. Dotted bars designate killing of VAC/*env*-infected target cells; solid bars designate killing of the paired VAC/*lac*-infected control target cells. All values are given at a representative E:T ratio of 50:1. HLA antigens shared by effector and target cells are noted in each case. Standard error bars are shown for experiments performed multiple times. HLA serotypes of the subjects whose cells were used in these experiments were: Subject A, HLA-A1, 29; B13, 35; Cw1, w4; DR1, 5; DRw52; DQw1, 3; subject B, HLA-A1, —; B51, —; C, —; DR2, 5; DRw52; DQw1, 3; subject C, HLA-A1, 11; B8, w62; Cw4, —; DR3, —; DRw52; DQw2; subject D, HLA-A2, 3; B35, 44; Cw4, w5; DR1, 2; DQw1; subject E, HLA-A2, 3; Bw57, w60; Cw3, w6; DR7, w9; DRw53; DQw2, 3; subject F, HLA-A2, 3; B7, w55; Cw3, w7; DR2, w6; DRw52, DQw1.



These results demonstrate the presence of HIV-specific cytotoxic T cells in persons infected with HIV. Despite the likelihood that the eight seropositive subjects are infected with heterologous HIV genotypes^{26,27}, all exhibited an HIV *env*-specific cytotoxic response against a single expressed *env* gene, suggesting that the cytotoxic effectors recognize shared envelope epitopes. A *gag*-specific cytotoxic response was also present, but to a lesser degree. Although the *gag* gene product has not yet been shown to be on the cell surface, there is precedent in other viral systems for internal viral structural proteins eliciting cytotoxic responses^{11,14,28,29}.

A feature of HIV infection that remains unexplained is why some individuals progress to ARC and AIDS, whereas others remain healthy, despite frequent isolation of virus from their blood. Whether the cellular immune response to HIV described here correlates with disease stage or contributes to protection from disease progression is unclear. In that results obtained from this assay system are highly reproducible, longitudinal analysis of HIV-specific cytotoxicity in individuals with disease progression is possible. By constructing HIV-vaccinia recombinant vectors with mutations in the *env* and *gag* gene inserts, it may also be possible to characterize the viral targets of this cytotoxic response further. This assay should also prove useful in evaluating AIDS vaccine strategies by providing a means of measuring cellular immune responses to candidate vaccines.

We thank T. Fuller, D. Fitzpatrick and J. Burroughs for assistance with HLA typing, R. C. Gallo, P. Earl and F. Preffer for materials and advice, M. Bahnam for invaluable technical assistance and J. Steele for preparation of this manuscript. B.D.W. is supported by a grant from the American Cancer Society. R.T.S. and M.S.H. are supported by grants from the NIH.

Received 2 April; accepted 29 May 1987

1. Zinkernagel, R. M. & Welsh, R. M. *J. Immun.* **117**, 1495-1502 (1976).
2. Yap, K. L., Ada, G. L. & McKenzie, I. F. C. *Nature* **273**, 238-239 (1978).
3. Lin, Y.-L. & Askonas, B. A. *J. exp. Med.* **154**, 225-234 (1981).
4. Quinnan, G. V. *et al. New Engl. J. Med.* **307**, 6-13 (1982).
5. McMichael, A. J., Gotch, F. M., Noble, G. R. & Beare, P. A. *S. New Engl. J. Med.* **309**, 13-17 (1983).
6. Townsend, A. R. M. & McMichael, A. J. *Prog. Allergy* **36**, 10-43 (1985).
7. Flyer, D. C., Burakoff, S. J. & Faller, D. V. *Nature* **305**, 815-818 (1983).
8. Kannagi, M. *et al. J. Immun.* **130**, 2942-2946 (1983).
9. Mitsuya, H. *et al. J. exp. Med.* **158**, 994-999 (1983).
10. Van der Hoorn, F. A., Lahaye, T., Müller, V., Ogle, M. A. & Enders, H. D. *J. exp. Med.* **162**, 128-144 (1985).
11. Holt, C. A., Osorio, K. & Lilly, F. *J. exp. Med.* **164**, 211-226 (1986).
12. Smith, G. L., Murphy, B. R. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7155-7159 (1983).
13. Bennick, J. R., Yewdell, J. W., Smith, G. L., Moller, C. & Moss, B. *Nature* **311**, 578-579 (1984).
14. Yewdell, J. W., Bennick, J. R., Smith, G. L. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1785-1789 (1985).
15. Bennick, J. R., Yewdell, J. W., Smith, G. L. & Moss, B. *J. Virol.* **57**, 786-791 (1986).
16. McMichael, A. J., Michie, C. A., Gotch, F. M., Smith, G. L. & Moss, B. *J. gen. Virol.* **67**, 719-726 (1986).
17. Zarling, J. M., Moran, P. A., Laskey, L. A. & Moss, B. *J. Virol.* **59**, 506-509 (1986).
18. Earl, P. E. *et al. Science* **234**, 728-731 (1986).
19. Chakrabarti, S., Robert-Guroff, M., Wong-Staal, F., Gallo, R. C. & Moss, B. *Nature* **320**, 535-537 (1986).
20. Hu, S. L., Kosowski, S. G. & Dalrymple, J. M. *Nature* **320**, 537-540 (1986).
21. Lifson, J. D. *et al. Nature* **323**, 725-728 (1986).
22. Ratner, L. *et al. Nature* **313**, 277-284 (1985).
23. Salahuddin, S. Z. *et al. Virology* **129**, 51-64 (1983).
24. Blumberg, R. C. *et al. J. infect. Dis.* **155**, 877-889 (1987).
25. Perrin, L. H., Zinkernagel, R. M. & Oldstone, M. B. A. *J. exp. Med.* **146**, 949-969 (1977).
26. Wong-Staal, F. *et al. Science* **229**, 759-762 (1985).
27. Starcich, B. R. *et al. Cell* **45**, 637-648 (1986).
28. Townsend, A. R. M., McMichael A. J., Carter, N. P., Huddleston, J. A. & Brownlee, G. G. *Cell* **39**, 13-25 (1984).
29. Yewdell, J. W. *et al. J. exp. Med.* **163**, 1529-1538 (1986).
30. Ho, D. D. *et al. Science* **226**, 451-453 (1984).
31. Essex, M. *et al. Science* **220**, 859-862 (1983).
32. Lazarovits, A. I., Kurnick, J. T., Camerini, D., Karsh, J. & Colvin, R. B. in *Leucocyte Typing III*, Proc. 3rd Int. Workshop on Leucocyte Differentiation Antigens (ed. McMichael A.) (Oxford Univ. Press, in the press).

AIDS virus-specific cytotoxic T lymphocytes in lung disorders

Fernando Plata*, Brigitte Autran†, Livia Pedroza Martins*, Simon Wain-Hobson*, Martine Raphaël‡, Charles Mayaud‡, Michel Denis‡, Jean-Marc Guillon‡ & Patrice Debré†

* Laboratoire de Biologie et d'Immunologie Moléculaires des Rétrovirus, Institut Pasteur, 75724 Paris Cédex 15, France

† Laboratoire d'Immunologie Cellulaire et Tissulaire, Hôpital Pitié-Salpêtrière, 75013 Paris, France

‡ Service de Pneumologie, Hôpital Tenon, 75010 Paris, France

Human immunodeficiency virus (HIV) is implicated in the development of AIDS (acquired immune deficiency syndrome)^{1,2}. HIV infection leads to the generation of HIV-specific thymus-derived (T) lymphocytes in humans^{3,4} and apes⁵. We describe an experimental system permitting the quantitative and systematic analysis of HIV-specific cytotoxic T lymphocytes (CTL). Functional, HIV-specific CTL are obtained by broncho-alveolar lavage (BAL) from the lungs of seropositive patients with lymphocytic alveolitis⁶. These alveolar CTL: (1) recognize and kill HIV-infected alveolar macrophages *in vitro* under autologous, but not heterologous, conditions; (2) correspond to standard CTL as they express the CD3 and CD8 surface markers, but not the CD4 marker; and (3) are restricted by class I HLA transplantation antigens in their cytotoxic activities. We propose the hypothesis that interactions between HIV-specific CTL and infected macrophages induce major inflammatory reactions in seropositive patients.

Analyses⁶⁻¹⁰ of BAL fluids or lung biopsies from HIV-positive patients suffering from lymphocytic alveolitis and lymphocytic

interstitial pneumonitis (LIP) have consistently indicated an abnormally high infiltration of CD8-positive T lymphocytes. Moreover, infectious HIV has been detected in BAL supernatants¹¹, lung biopsies^{6,7} and BAL leukocytes^{11,12} from these patients. Immuno-enzymatic assays performed in our laboratory with a monoclonal antibody specific for HIV protein p18 (ref. 13) show that 12-58% of the macrophages in BAL fluids are positive for HIV (data not shown). A characteristic of these patients is that they have not developed viral, bacterial or fungal lung infections^{6,11}.

We have studied the BAL fluids from a series of HIV-positive patients with lymphocytic alveolitis. The cells recovered consisted of 20-77% macrophages, 17-75% lymphocytes and 1-7% polymorphonuclear cells. Analysis by flow cytometry (data not shown) indicated that the lymphocyte subpopulation consisted of 80-93% T cells bearing the CD3 surface marker. A minority of these cells (3-13%) were CD4-positive helper T lymphocytes. The remaining 65-87% were double positives for the CD8 and D44 (ref. 14) markers which identify the cytotoxic T lymphocyte subset.

The adherent macrophage subpopulation in BAL fluids was easily separated from the non-adherent lymphocyte subpopulation by incubation in plastic tissue culture flasks. Figure 1a shows a Southern blot of alveolar macrophage DNA from patient B.N. When hybridized with a DNA probe containing the complete HIV genome, specific bands corresponding to 9.6 kilobases (uncut DNA) and 2.0 kilobases (*Hind*III-digested DNA) were detectable, confirming the presence of HIV in these cells.

A separate sample of the same macrophages from patient B.N. was labelled with ⁵¹Cr and used as target cells to alveolar lymphocytes in an 18-h cytotoxicity assay (Fig. 1b). Heterologous HIV-positive alveolar macrophages from patient D.S. were also used as target cells. Figure 1b shows that B.N. alveolar lymphocytes, in the absence of previous *in vitro* culture or stimulation, efficiently killed B.N. macrophages at low lym-

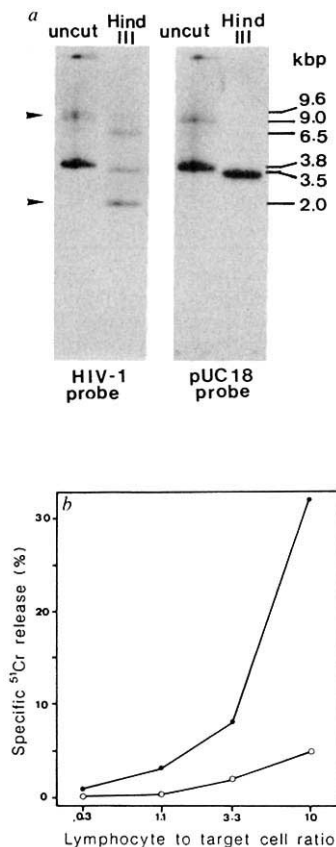


Fig. 1 Alveolar macrophages from patient B.N. contained HIV DNA and were killed by autologous alveolar lymphocytes. *a*, The Southern blot of B.N. macrophage DNA was hybridized with a ³²P-labelled probe containing the complete HIV-1 genome or with a control plasmid DNA probe (pUC18). The arrow heads indicate B.N. DNA fragments (2.0 and 9.6 kb) which hybridized specifically to HIV DNA. Bacterial DNA was also detected, probably corresponding to bacteria phagocytosed by these macrophages. *b*, Cytotoxicity mediated by B.N. alveolar lymphocytes was determined on ⁵¹Cr-labelled autologous alveolar macrophages (●) in an 18-h assay. Heterologous macrophages from patient D.S. (○) were also used as target cells to test for specificity of lysis. Spontaneous ⁵¹Cr release values were 21% (B.N. targets) and 28% (D.S. targets) of total incorporated radioactivity. Kbp, Kilobase-pair. **Methods.** BAL leukocytes were incubated in plastic tissue culture flasks at 10⁶ cells per ml of RPMI-1640 culture medium supplemented with 20% fetal bovine serum and antibiotics. After 2 h incubation at 37 °C, non-adherent cells were recovered by repeated pipetting of the attached macrophage monolayer with RPMI culture medium. Macrophages were recovered after 5 min incubation in 0.05% trypsin and 0.02% EDTA at 37 °C. After counting, 2 × 10⁷ macrophages from patient B.N. were submitted to DNA extraction according to Maniatis²⁸. A separate sample of 2 × 10⁶ macrophages was labelled with 300 μCi Na₂ ⁵¹CrO₄ as described elsewhere²³. Untreated or *Hind*III-digested DNA (5 μg aliquots) was submitted to electrophoresis in a 0.8% agarose gel, and then transferred to nitro-cellulose filter²⁸. The filter was first hybridized to a ³²P-labelled 9-kb probe containing the complete HIV-I genome¹⁷, washed and exposed to autoradiography. The same filter was then dehybridized²⁸ and rehybridized with a ³²P-labelled pUC18 probe²² containing only plasmid DNA. Fragment sizes were estimated according to migration of phage λ *Hind*III DNA fragments applied to the same gel. Cytotoxicity assays were performed with 10⁴ ⁵¹Cr-labelled alveolar macrophages distributed in flat-bottomed 96-well microtitre plates. Alveolar lymphocytes from patient B.N. were added to each well at different concentrations in a final volume of 200 μl per well. Plates were incubated for 18 h at 37 °C, and percent specific ⁵¹Cr release was determined from each well supernatant as described previously²³.

Table 1 Specificity of effector lymphocytes

Experiment	Lymphocyte donor	Target cells	Expression of		Lymphocyte-to-target cell ratio		
			HLA-A2	HIV antigen(s)	10:1	3:1	1:1
1	L.T.	L.T. macrophages	+	+	7	2	0
		B.S. macrophages	-	+	2	0	0
		P815-A2	+	-	-0.5	0	-0.8
		P815-A2-env-LAV	+	+	11	2	1
2	G.S.	G.S. macrophages	+	+	15	7	4
		V.E. macrophages	-	+	-1.5	1.2	-0.3
		P815-A2	+	-	-1.1	-1.4	-0.4
		P815-env-LAV	-	+	-3.8	-0.6	0.5
		P815-A2-env-LAV	+	+	27	9	3
3	D.S.	D.S. macrophages	+	+	18	6	1
		L.T. macrophages	+	+	10	4	2
		V.E. macrophages	-	+	-0.9	-0.8	1
		P815-A2	+	-	1.6	1.5	-1.9
		P815-env-LAV	-	+	0.4	0	-0.3
		P815-A2-env-LAV	+	+	15	10	2

Alveolar leukocytes from patients L.T., G.S., D.S., B.S. and V.E. were separated into lymphocyte and macrophage fractions by adherence to plastic (see Fig. 1). Alveolar lymphocytes were tested for cytotoxicity on ⁵¹Cr-labelled macrophages or transfected P815 cells in 18-h assays. Mouse P815-A2 cells contain and express the human HLA-A2 gene¹⁵. P815-env-LAV and P815-A2-env-LAV cells were obtained by CaPO₄ transfection of P815 or P815-A2 cells, respectively, with the 3.1-kilobase (kb) *Sal*I-*Xho*I DNA fragment from HIV-1 (BRU)¹⁷ containing the *env* gene, placed under control of its own LTR promoter, and inserted in plasmid pUC18²². A non-coding 521-base pair (bp) *Ava*I-*Bgl*II fragment from hepatitis B virus was inserted 3' to the HIV *env* gene to provide for poly-A, splicing acceptor sequences and other termination sequences. HIV *env* gene DNA was co-transfected²³ with plasmid pAG-60 (ref. 24), which contains the bacterial gene for resistance to antibiotic G418. Transfectants were selected in culture medium containing 300 μg ml⁻¹ G418 (Gibco Europe Ltd, Scotland), cloned by limiting dilution, and screened by immunofluorescence²³ with sera from AIDS patients and with monoclonal antibody KL 39.19 to *env* gene-coded gp 150 (a gift from Dr Luc Montagnier, Institut Pasteur). HLA-A2 positivity of L.T., G.S. and D.S. patients was determined serologically with peripheral blood lymphocytes in a standard complement-mediated cytotoxicity assay using HLA-A2-specific antisera. HIV infection of the macrophage target cells was confirmed by immunoenzymatic labelling using CVK-A1 anti-p18 monoclonal antibodies¹³.

Table 2 Surface markers of alveolar effector lymphocytes

Experiment	Lymphocyte donor	Treatment of effector lymphocytes	Target cells	E:T ratio	None	Specificity of monoclonal antibodies			
						CD3	CD4	CD8	D44
						% Cytotoxicity (% decrease)			
1	D.S. (HLA-A2)	Blocking antibodies	D.S. macrophages	5:1	8	3 (60)	11	3 (60)	-1 (100)
			P815-A2-env-LAV	5:1	13	9 (31)	17	-0.5 (100)	2 (85)
2	B.N.	C' depletion	B.N. macrophages	10:1	32	-0.4 (100)	30 (6)	7 (78)	22 (31)
3	C.T.	C' depletion	C.T. macrophages	5:1	49	11 (78)	35 (28)	2 (96)	4 (92)

The identity of effector alveolar lymphocytes was determined either by blocking of effector functions (patient D.S.) or by rabbit complement (C') depletion of effector lymphocytes (patients B.N. and C.T.) using cytotoxic monoclonal antibodies specific for CD3 (OKT3; ref. 25), CD4 (OKT4, ref. 25; and IOT4, Immunotech S.A.), CD8 (B9.84; ref. 26) and D44 (ref. 14) lymphocyte surface marker molecules. Blocking experiments were performed at a lymphocyte-to-macrophage ratio of 5:1 as described previously²⁵ with a final antibody dilution of 1/20 (OKT3 culture supernatant) or 1/120 (OKT4, B9.84 and D44 ascites fluids). C' depletion was performed with 0.5 to 1×10^6 alveolar lymphocytes incubated with the same antibodies diluted 1/10 (OKT3), 1/100 (IOT4 and B9.84) and 1/200 (D44) and with rabbit complement (1/18), as described elsewhere²⁷. The remaining lymphocytes were adjusted to a ratio of 5 or 10 lymphocytes to one ^{51}Cr -labelled target cell, as indicated, and assayed for cytotoxicity in a 18-h assay.

phocyte-to-target cell ratios and failed to kill heterologous macrophages from patient D.S. Cytotoxic alveolar lymphocytes were similarly detected in six other HIV-positive patients with lymphocytic alveolitis when tested on HIV-infected autologous macrophages (Fig. 2, Tables 1 and 2, and data not shown).

Figure 2 summarizes the cytotoxicity data obtained with three other HIV-infected patients with lymphocytic alveolitis. Patients L.T. (Fig. 2a) and C.T. (Fig. 2c) showed cytotoxicity patterns comparable to those obtained with patient B.N. (Fig. 1b), as L.T. and C.T. alveolar lymphocytes killed autologous macrophages more efficiently than heterologous HIV-positive macrophages. In contrast, lymphocytes from patient G.S. (Fig. 2b) killed heterologous macrophages from patient L.T. more efficiently than autologous (G.S.) macrophages (24 compared to 11% lysis, at a lymphocyte-to-macrophage ratio of 84:1). Assuming that cytotoxic alveolar lymphocytes are HLA-restricted CTL, the striking cross reactivity observed between G.S. lymphocytes and L.T. macrophages could be ascribed to a shared HLA specificity as patients L.T. and G.S. were both positive for HLA-A2 antigen (data not shown).

We tested this assumption by constructing a mouse target cell which expressed both HLA-A2 and HIV envelope proteins, and which was easily lysed by human CTL. P815-A2 cells¹⁵ (previously transfected with plasmid DNA containing the HLA-A2 gene) were additionally transfected with the HIV *env* gene. The stable cell line 'P815-A2-env-LAV' was thus derived. Mouse cell lines expressing HLA-A2 only (P815-A2; ref. 15) and the HIV

env gene only (P815-env-LAV) were used as controls. Table 1 shows that alveolar lymphocytes from HLA-A2-positive patients L.T., G.S. and D.S. killed autologous HIV-infected macrophages as well as P815-A2-env-LAV mouse cells. P815-A2 and P815-env-LAV cells were not killed, indicating that joint expression of HLA-A2 and HIV envelope protein was required. Table 1 also shows that heterologous HIV-infected macrophages were killed provided they shared the HLA-A2 specificity (Patient D.S., experiment 3). Thus, a proportion of the alveolar lymphocytes tested in Table 1 were restricted by class I HLA-A2 transplantation antigen. Moreover, these data also indicate that HIV envelope glycoprotein is detected by alveolar lymphocytes on the surface of the infected macrophages.

Table 2 shows that the cytotoxic alveolar lymphocytes detected in our assay system were T lymphocytes positive for the CD3, CD8 and D44 surface markers and negative for CD4. These conclusions were drawn from the blocking activity of CD3-, CD8- and D44-specific monoclonal antibodies (experiment 1), and from the depletion of cytotoxic effector cells following incubation of alveolar lymphocytes with the same antibodies and rabbit complement (experiments 2 and 3). When tested under comparable conditions, anti-CD4 monoclonal antibodies had no effect. These lymphocytes thus belong to the classic cytotoxic subset of T lymphocytes.

In view of the extreme genetic variability of HIV isolates¹⁶ it is probable that our patients were infected by different HIV strains. Despite this, CTL from different patients positive for HLA-A2 killed heterologous HIV-infected macrophages provided they expressed HLA-A2 transplantation antigen (Fig. 2 and Table 1). Furthermore, these HLA-A2-positive CTL also killed P815-A2-env-LAV cells which expressed the *env* gene of the first HIV isolate¹⁷ (LAV-1-BRU). Conserved, cross-reactive epitopes among HIV antigens are thus target antigens for HIV-specific CTL in humans.

The alveolar CTL activities described here directly reflect the status of these lymphocytes as they exist *in vivo*, since they were tested immediately after BAL in the absence of *in vitro* culture or stimulation. Cytotoxic activities at high lymphocyte-to-macrophage ratios rarely surpass 30% lysis, indicating either a relative resistance of alveolar macrophages to CTL attack or an inadequate presentation of HIV antigens by the targets. It is equally intriguing that P815-A2-env-LAV target cells were not killed by alveolar CTL at levels higher than those observed in Tables 1 and 2. These transfected cells uniformly expressed both HLA-A2 and HIV envelope glycoprotein antigens, and were killed with high efficiencies by HIV-specific CTL lines derived from peripheral blood lymphocytes (data not shown). One possible explanation for the relatively low lysis of P815-A2-env-LAV target cells by alveolar CTL is that a fraction of these CTL is directed to HIV-induced antigens different from the envelope

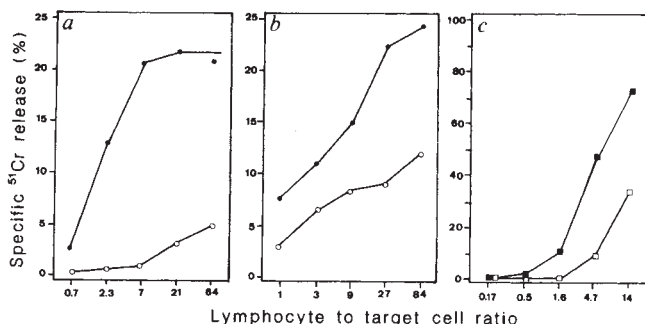


Fig. 2 Cytotoxic activities of alveolar lymphocytes from three different patients. Alveolar lymphocytes from patients L.T. (a) and G.S. (b) were assayed for cytotoxicity on ^{51}Cr -labelled alveolar macrophages from patients L.T. (●) and G.S. (○). Alveolar lymphocytes from patient C.T. (c) were tested on alveolar macrophages from patients C.T. (■) and D.S. (□). Cytotoxicity assays were performed at various lymphocyte-to-target cell ratios for 18 h. Spontaneous ^{51}Cr release values varied between 21% and 32% of total incorporated label.

antigen (for example, products of the *gag* gene) and consequently cannot be detected with these target cells. In this same context, we observed that the intensity of CTL responses varied widely among different patients, as well as within single patients studied at different time points (patients L.T. and C.S., Fig. 2 and Table 1). Clinical correlations remain to be drawn.

The question remains as to the beneficial or deleterious effect of HIV-specific CTL activities in seropositive patients. Our results, in conjunction with previous observations by others^{6,8-10}, suggest that CD8-positive T lymphocytes with cytotoxic activities induce local inflammations in the lungs by their interactions with HIV-infected alveolar macrophages. A similar hypothesis has been formulated for another lung disorder, active pulmonary sarcoidosis, where lymphocytic alveolitis is also observed^{18,19}. The HIV-infected macrophages would thus serve not only as a reservoir for the virus^{6,7,11,12}, but also as an inflammatory agent opening the way for bacterial or fungal infections of the lung. Similar considerations can be formulated for other organs including the brain^{20,21}, where local inflammatory reactions could eventually be induced in the areas surrounding HIV-infected macrophages.

We thank Dr Marie-Rose Marescot for determining the HLA phenotypes of our patients, Drs Luc Montagnier and David Klatzman for their gifts of monoclonal antibodies KL 39.19 and CVK-A1, respectively, Drs Elisabeth Gomard and Jean-Paul Lévy for the gift of P815-A2 cells and helpful advice, Dr Philippe Kourilsky for encouragement and help in performing this work,

and Ms Michèle Chahine for typing the manuscript. This project was supported by grants from Institut Pasteur, Institut National de la Santé et de la Recherche Médicale, and Ministère de la Recherche et de l'Industrie, France.

Received 16 April; accepted 6 June 1987.

1. Barré-Sinoussi, F. *et al. Science* **220**, 868-871 (1983).
2. Gallo, R. C. *et al. Science* **220**, 865-867 (1983).
3. Walker, C. M. *et al. Science* **234**, 1563-1566 (1986).
4. Zagury, D. *et al. Nature* **326**, 249-250 (1987).
5. Zarling, J. M. *et al. Nature* **323**, 344-346 (1986).
6. Chayt, K. J. *et al. J. Am. Med. Assn* **256**, 2356-2359 (1986).
7. Gartner, S. *et al. Science* **233**, 215-219 (1986).
8. Wallace, J. M. *et al. Am. Rev. Respir. Dis.* **130**, 786-790 (1984).
9. Young, K. R. *et al. Ann. intern. Med.* **103**, 522-533 (1985).
10. Venet, A. *et al. Bull. Eur. Physiopathol. Respir.* **21**, 535-543 (1985).
11. Ziza, J. M. *et al. N. Engl. J. Med.* **313**, 183 (1985).
12. Salahuddin, S. Z. *et al. Blood* **68**, 281-284 (1986).
13. Chassagne, J. *et al. J. Immun.* **136**, 1442-1445 (1986).
14. Calvo, C. F. *et al. J. Immun.* **132**, 2345-2349 (1984).
15. Gomard, E. *et al. Nature* **319**, 153-154 (1986).
16. Wong-Staal, F. & Gallo, R. C. *Nature* **317**, 395-403 (1985).
17. Alizon, M. *et al. Nature* **312**, 757-760 (1984).
18. Hunninghake, G. W. *Am. Rev. Respir. Dis.* **129**, 569-572 (1984).
19. Pinkston, P. *et al. N. Eng. J. Med.* **308**, 793-800 (1983).
20. Shaw, G. M. *et al. Science* **227**, 177-182 (1985).
21. Gartner, S. *et al. J. Am. Med. Assn* **256**, 2365-2371 (1986).
22. Vieira, J. & Messing, J. *Gene* **19**, 259-268 (1982).
23. Plata, F. *et al. Cell* **48**, 231-240 (1987).
24. Colbère-Garapin, F. *et al. J. molec. Biol.* **150**, 1-14 (1981).
25. Reinherz, E. L. & Schlossman, S. F. *Cell* **19**, 821-827 (1980).
26. Malissen, B. *et al. Eur. J. Immun.* **12**, 739-747 (1982).
27. Plata, F. *J. Immun.* **134**, 1312-1319 (1985).
28. Maniatis, T. *et al. Molecular Cloning: a Laboratory Manual* (Cold Spring Harbor Laboratory Publications, Cold Spring Harbor, 1982).

Expression and function of CD4 in a murine T-cell hybridoma

Barry P. Sleckman, Andrew Peterson*, William K. Jones, Judith A. Foran, Julia L. Greenstein, Brian Seed* & Steven J. Burakoff

Division of Pediatric Oncology, Dana-Farber Cancer Institute, Departments of Pediatrics and Pathology, Harvard Medical School, Boston, Massachusetts 02115, USA

* Department of Molecular Biology, Massachusetts General Hospital, Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA

The CD4 (T4) antigen was originally described as a phenotypic marker specific for helper T cells, and has recently been shown to be the receptor for the human immunodeficiency virus (HIV)¹⁻⁴. Functional studies using monoclonal antibodies directed at CD4 and major histocompatibility complex (MHC) class II molecules led to the suggestion that CD4 binds to the MHC class II molecules expressed on stimulator cells, enhancing T-cell responsiveness by increasing the avidity of T cell-stimulator cell interaction and/or by transmitting a positive intracellular signal⁵⁻¹¹. But recent evidence that antibodies to CD4 inhibit T-cell responsiveness in the absence of any putative ligand for CD4 has been interpreted as suggesting that antibody-mediated inhibition may involve the transmission of a negative signal via the CD4 molecule instead¹²⁻¹⁴. We have infected a murine T-cell hybridoma that produces interleukin 2 (IL-2) in response to human class II HLA-DR antigens with a retroviral vector containing CD4 cDNA. The resulting CD4-expressing hybridoma cell lines produce 6- to 20-fold more IL-2 in response to HLA-DR antigens than control cell lines. Furthermore, when antigen levels are suboptimal, the response of the cell lines is entirely CD4-dependent. The data presented here clearly demonstrate that CD4 can enhance T-cell responsiveness and may be crucial in the response to suboptimal levels of antigen.

A murine T-cell hybridoma (By155.16) that produces IL-2 in response to stimulator cells (JY and Daudi) expressing HLA-DR antigens was infected with the defective retroviruses MNST4

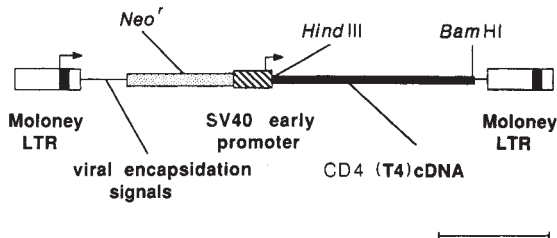


Fig. 1 Structure of the retroviral vector MNST4. MNST4 has retroviral transcription and integration sequences in the LTRs derived from Moloney murine leukaemia virus. The neomycin-resistant element is the APH (3') II gene from Tn5 and confers resistance to G418 in eukaryotic cells. The messenger RNA for the *neo* gene and the viral genomic RNA are produced by transcription initiation in the left LTR and polyadenylation in the right LTR. The CD4 cDNA is expressed from a subgenomic RNA initiated at an internal promoter, the SV40 early promoter. The xenotropic helper cell line, MXEN (A. Peterson, unpublished data), was transfected with pMNST4 and pMN (pMNST4 plasmid without the CD4 cDNA insert) using a modification of the DEAE-dextran procedure of ref. 24 (B. Seed and A. Aruffo, unpublished). Forty-eight hours after transfection the supernatants from the cells were passed through a 0.45 μ m filter and were used to infect the amphotropic helper line, DAMP (A. Peterson, unpublished), in the presence of 10 μ g ml⁻¹ polybrene. Selection was initiated 48 h after infection using 500 μ g ml⁻¹ G418. Pooled colonies of G418^r DAMP cells were used as the MNST4 DAMP and the MN DAMP producer lines. Scale bar, 1.0 kilobase pairs.

and MN produced by the MNST4 DAMP and MN DAMP cell lines, respectively (see Fig. 1 legend). The defective retrovirus produced by MNST4 DAMP expresses the *neo* gene driven by the Moloney leukaemia virus long terminal repeat (MLV LTR) and expresses the complementary DNA encoding human CD4 driven by the simian virus 40 (SV40) promoter as shown in Fig. 1. The defective retrovirus produced by MN DAMP expresses only the *neo* gene driven by the MLV LTR. As a provirus, the MNST4 genome imparts resistance to the antibiotic G418 and stable long-term expression of CD4 to the host cell, whereas the MN proviral genome imparts only resistance to G418.

Table 1 Enhanced responsiveness of CD4-expressing murine T-cell hybridomas

	Daudi cells	Antibody added	Production of IL-2 (U ml ⁻¹)		
			16M-9 (CD4 ⁻)	16T4-3 (CD4 ⁺)	16T4-9 (CD4 ⁺)
Expt 1	-	None	<10	<10	<10
	+	None	275	1,570	5,060
	+	OKT4	298	780	2,532
	+	OKT4F	275	243	885
	+	Leu3a	276	30	400
Expt 2	-	None	<10	ND	<10
	+	None	239	ND	4,301
	+	GK1.5	204	ND	3,772
	+	Leu3a	310	ND	153
	+	TS1/16	30	ND	24

T-cell hybridoma cells (5×10^4) were incubated with 3×10^5 Daudi cells in a total volume of 1.2 ml DMEM-10 containing no antibody, 500 ng ml^{-1} OKT4 (CD4), 500 ng ml^{-1} OKT4F (CD4), 360 ng ml^{-1} Leu3a (CD4), $1,000 \text{ ng ml}^{-1}$ GK1.5 (L3T4), or a 1:1,000 dilution of an ascites containing the monoclonal antibody TS1/16 (HLA-DR). Cultures were incubated for 24 h at 37°C in 5% CO_2 . Supernatants were harvested, titrated by serial twofold dilutions and assayed for IL-2 by their ability to support the proliferation of the IL-2-dependent line, CTLL20 (refs 22, 23). Proliferation was assessed by the incorporation of ^3H -thymidine in a 4-h pulse after a 20-h culture. Experiments 1 and 2 were each carried out at least three times. Single representative experiments are given. Data are presented as units per ml relative to the ^3H -thymidine incorporation of CTLL20 from a standard rat concanavalin A (Con A) supernatant. The half-maximal incorporation from the Con A titration was defined as 100 units IL-2 per ml.

ND, not determined.

Two infections of By155.16 with MNST4 DAMP were carried out, for which 33 G418 resistant lines were isolated. Twenty-five of these lines expressed CD4, as determined by flow cytometric analysis, of which 18 were responsive to the HLA-DR-expressing Daudi cell line. The unresponsive lines were unresponsive to both the Daudi and the JY cell lines, which both express HLA-DR antigens. This unresponsiveness may have been induced by the selection process or the integration of the retroviral genome into the host genome, resulting in the inability of these cells to express certain genes required for the production of IL-2 in response to antigen. Of the functional CD4-expressing lines, the least responsive (16T4-3) and the most responsive (16T4-9) lines were chosen for further functional studies. A single infection of By155.16 with MN DAMP was also carried out, from which 9 G418-resistant lines were isolated. Upon functional analysis the MN DAMP-infected cell lines were found to produce at most twice as much IL-2 in response to the Daudi cell line, as compared to the parent line By155.16. The most responsive of these lines (16M-9) was chosen for further functional studies. All cell lines chosen for functional studies, and the parent line By155.16, responded to the Daudi cell line, which expresses only MHC class II and no MHC class I antigens. Furthermore, this response is entirely blocked by the HLA-DR-specific monoclonal antibodies TS1/16 (ref 15; Table 1) and LB3.1 (ref. 16; data not shown). All cell lines studied respond to purified HLA-DR antigens in liposomes, confirming the specificity of these cells for HLA-DR (data not shown). Figure 2 shows a flow cytometric analysis of the 16M-9, 16T4-3, and 16T4-9 cell lines using monoclonal antibodies to CD4 (OKT4)¹, the murine T-cell receptor (TCR) (F23.1)¹⁷ and murine CD4 (L3T4; GK1.5)¹⁸. Human CD4 expression by 16T4-9 was about tenfold less than that of human peripheral blood lymphocytes.

The CD4-expressing cell lines demonstrate a 6- to 20-fold increased IL-2 response to 3×10^5 stimulator cells as compared to the non-CD4-expressing 16M-9 line (Table 1). It should be noted that the less responsive line, 16T4-3, exhibits a lower level of expression of CD4 and TCR.

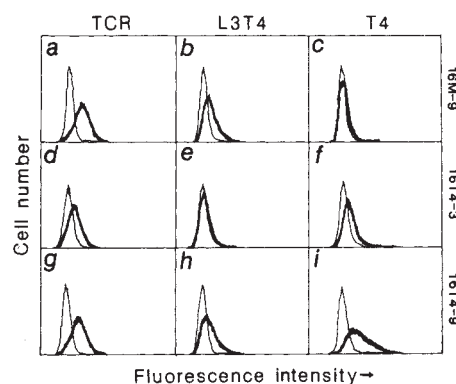


Fig. 2 Flow cytometric analysis of the murine TCR, murine CD4 (L3T4), and human CD4 (T4). Lines 16M-9, 16T4-3 and 16T4-9 were stained with OKT4 (CD4) GK1.5 (L3T4) and F23.1 (TCR) monoclonal antibodies. By155.16 was produced by fusion of an *in vivo* JY (human B-cell line) primed and *in vitro* boosted C57BL/6 spleen cell population with the HAT-sensitive thymoma BW5147. The cells were fused using a standard protocol and selected in HAT-containing medium. The hybridomas were screened for production of IL-2 in the presence of JY. By155 was found to be antigen specific and was subsequently cloned by limiting dilution. By155.16 was infected with MNST4 DAMP on two occasions by incubating 2×10^5 cells in 10 ml of supernatant from a confluent plate of MNST4 DAMP cells with $10 \mu\text{g ml}^{-1}$ polybrene. By155.16 was infected with MN DAMP once in the same manner. Incubations were carried out for 48 h and cells were transferred to selection medium containing 2.0 mg ml^{-1} of the antibiotic G418 (Gibco). Neomycin-resistant lines were isolated from the MNST4 DAMP and the MN DAMP infections, giving 33 and 9 lines respectively, and analysed for expression of human CD4 (T4), murine CD4 (L3T4) and murine TCR. Cells were washed with phosphate-buffered saline (PBS) with 2% fetal calf serum and 0.02% azide and incubated in $50 \mu\text{l}$ of OKT4, GK1.5, or F23.1 antibody solution at 4°C for 30 min. Cells were then washed and incubated in fluorescein isothiocyanate (FITC)-conjugated F(ab')_2 anti-mouse immunoglobulin (OKT4 and F23.1) or anti-rat immunoglobulin (GK1.5) (Tago) before analysis using a fluorescence-activated cell sorter (Coulter). Ten thousand cells were analysed and dead cells were gated out of the analysis. Thin line, FITC-conjugated anti-mouse immunoglobulin or anti-rat immunoglobulin alone; thick line, test antibody plus FITC-conjugated anti-mouse immunoglobulin or anti-rat immunoglobulin.

The functional role of CD4 in this murine T-cell hybridoma was also examined by antibody blocking. The monoclonal antibodies OKT4, OKT4F and Leu3a, which recognize distinct epitopes on the CD4 molecule¹⁹, specifically decreased the responsiveness of the CD4-expressing cell lines (Table 1). The response of the CD4 infectants could be blocked to near that of 16M-9 with 500 ng ml^{-1} OKT4F or 360 ng ml^{-1} Leu3a (Table 1). This level of blocking was maximal, as increasing the concentration of either OKT4F or Leu3a did not increase the inhibition of IL-2 production (data not shown). At 500 ng ml^{-1} OKT4 inhibited the response of the CD4-expressing lines by 50% (Table 1). This level of inhibition was also maximal, as increasing the OKT4 MAb concentration to $50 \mu\text{g ml}^{-1}$ did not increase the blocking. The inability of OKT4 to inhibit function as well as Leu3a and OKT4F is not due to decreased binding to CD4 as, under conditions of excess antibody, all three antibodies stain CD4-expressing lines with equal intensity. The differences observed in the ability of these three antibodies to inhibit the responsiveness of the CD4-expressing lines may reflect their different abilities to inhibit the CD4/class II interaction or to induce a negative signal.

As determined by flow cytometric analysis, 16M-9 and 16T4-9 express murine CD4 (L3T4) whereas 16T4-3 does not (Fig. 2). The data in Table 1 clearly show that $1,000 \text{ ng ml}^{-1}$ of GK1.5 (a monoclonal antibody to murine CD4)¹⁸ had a minimal effect

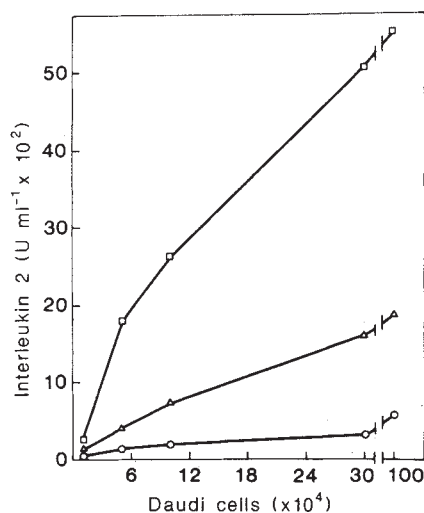


Fig. 3 Expression of CD4 increases response to low doses of Daudi. Assays were carried out as described in Table 1. Data shown are from a representative experiment of three experiments that were carried out: 16M-9 (○), 16T4-3 (△), 16T4-9 (□).

on the 16M-9 and 16T4-9 cell lines. Antibody titration experiments carried out on these cell lines in parallel with the CD4 (L3T4)-dependent T-cell hybridoma 3DT52.5¹⁰ showed that 1,000 ng ml⁻¹ of GK1.5 was more than tenfold in excess of the amount of antibody needed to inhibit the function of 3DT52.5 completely when stimulated with A20 (data not shown). These data suggest that if the murine CD4 expressed by the murine T-cell hybridomas is functional, it does not contribute substantially to the enhancement of responsiveness observed by the human CD4-expressing cell lines. The ability of antibodies to the human CD4 molecule to block the function of the CD4-expressing lines is not due to nonspecific steric effects of antibodies bound to the cell surface as antibodies to H-2 class I (M142)²⁰ and Thy 1.1 (22.1)²¹, both of which are expressed at higher densities than CD4 on all cell lines studied, had no significant effect on their function (data not shown). By155.16 has also been infected with a retroviral vector capable of imparting CD8 (T8) expression. The CD8-expressing lines isolated were found to produce approximately the same amount of IL-2 in response to the Daudi cell line as does 16M-9 (S. Ratnoffsky, *et al.*, manuscript in preparation). Thus, the enhanced responsiveness of the CD4-expressing lines is not simply due to an aberrant effect caused by the expression of foreign cell-surface protein in these cells.

It has been postulated that the increased avidity between the T cell and stimulator cell and/or the intracellular signal mediated by CD4 may be critical to T-cell responsiveness under conditions of suboptimal antigen. We investigated the effect of CD4 under conditions of suboptimal antigen by looking at the ability of the infectants to respond to limiting numbers of Daudi cells. In this case the level of antigen expressed on a single stimulator cell is not altered, but the probability that a stimulator cell will meet a T cell is decreased. Therefore, any alteration in the T cell that increases its ability to be activated should result in a greater response at lower levels of stimulator cells. Cell lines 16T4-3 and 16T4-9 respond better than 16M-9 to levels of Daudi cells ranging from 10⁴ to 10⁶; 16T4-9 produces as much IL-2 in response to 10⁴ Daudi cells as 16M-9 produces to 3 × 10⁵ Daudi cells (Fig. 3). These data suggest that the expression of CD4 by the parent line By155.16 allows a response to 30-fold fewer stimulator cells.

The data presented here demonstrate that CD4 is capable of enhancing T-cell responsiveness and may play a critical role in T-cell activation under conditions of suboptimal antigen stimulation.

This work was supported in part by grants from the NIH to S.J.B. and J.L.G. and a grant from Hoechst AG to B.S. and A.P. J.L.G. is the recipient of a Leukemia Society Special Fellowship. We thank Mary Crimmins for technical assistance, Gail Dudley for preparation of the manuscript and Dr Barbara Bierer for critical review of the manuscript. We also thank Dr Gideon Goldstein for providing us with the OKT4F antibody.

Received 28 May; accepted 2 June 1987.

1. Reinherz, E. L., Kung, P., Goldstein, G. & Schlossman, S. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4061-4065 (1979).
2. Dagleish, A. G. *et al. Nature* **312**, 763-767 (1984).
3. Klatzmann, D. *et al. Nature* **312**, 767-768 (1984).
4. McDougal, J. S. *et al. Science* **231**, 382-385 (1986).
5. Krensky, A. M., Reiss, C. S., Mier, J. W., Strominger, J. L. & Burakoff, S. J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2365-2369 (1982).
6. Meuer, S. C., Schlossman, S. F. & Reinherz, E. L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4395-4399 (1982).
7. Biddison, W. E., Rao, P. E., Talle, M. A., Goldstein, G. & Shaw, S. J. *exp. Med.* **156**, 1065-1076 (1982).
8. Marrack, P. *et al. J. exp. Med.* **158**, 1077-1091 (1983).
9. Greenstein, J. L., Malissen, B. & Burakoff, S. J. *J. exp. Med.* **162**, 369-374 (1985).
10. Greenstein, J. L., Kappler, J., Marrack, P. & Burakoff, S. J. *J. exp. Med.* **159**, 1213-1224 (1984).
11. Shaw, S., Goldstein, G., Springer, T. A. & Biddison, W. A. *J. Immun.* **134**, 3019-3026 (1985).
12. Bank, I. & Chess, L. *J. exp. Med.* **162**, 1294-1303 (1985).
13. Greenstein, J. L. & Burakoff, S. J. *Ann. Inst. Pasteur/Immun.* **138**, 134-138 (1987).
14. Tite, J. P., Sloan, A. & Janeway, C. A. *J. molec. cell. Immun.* **2**, 179-190 (1986).
15. Sanchez-Madrid, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 7489-7493 (1982).
16. Gorga, J. C., Knudsen, P. J., Foran, J. A., Strominger, J. L. & Burakoff, S. J. *Cell. Immun.* **103**, 160-173 (1986).
17. Staerz, U., Rammensee, H. G., Benedetto, J. D. & Bevan, M. J. *J. Immun.* **134**, 3994-4000 (1985).
18. Dialynas, D. P. *et al. Immun. Rev.* **74**, 29-56 (1983).
19. Sattentau, Q. J., Dagleish, A. G., Weiss, R. A. & Beverley, P. C. L. *Science* **234**, 1120-1123 (1986).
20. Mescher, M. F., Stallcup K. C., Sullivan, C. P., Turkewitz, A. P. & Hermann, S. H. *Meth. Enzym.* **92**, 86-109 (1983).
21. Rothstein, A. M. *et al. J. Immun.* **122**, 2491-2497 (1979).
22. Ely, J. M. *et al. J. Immun.* **127**, 2345-2349 (1981).
23. Gillis, S., Ferm, M. M., Ou, W. & Smith, K. A. *J. Immun.* **120**, 2027-2032 (1978).
24. Sussman, D. J. & Milman, G. *Molec. cell. Biol.* **4**, 1641-1643 (1984).

Genetic evidence that zinc is an essential co-factor in the DNA binding domain of GAL4 protein

Mark Johnston

Department of Genetics, Washington University School of Medicine, St Louis, Missouri 63110, USA

The 'cysteine-zinc DNA binding finger' is a recently identified sequence motif that is present in a wide variety of transcriptional regulatory proteins and is thought to be directly involved in DNA binding¹. It has been proposed that an essential component of this structure is a zinc ion bound between two pairs of cysteine residues. The *GAL4*-encoded protein of *Saccharomyces cerevisiae*, which binds to DNA and activates the transcription of several genes, contains this sequence motif. Here I describe a *gal4* mutant with an alteration in the cysteine-zinc DNA binding finger whose defect is corrected *in vivo* by high concentrations of ZnCl₂. The DNA binding activity of the altered protein from this mutant is restored by ZnCl₂ *in vitro*. This is evidence that the GAL4 protein indeed contains zinc ions essential for its DNA binding activity.

It has been proposed that the 'zinc binding finger' contains a zinc ion bound between two pairs of cysteine (or histidine) residues (see Fig. 1a). The amino acids between the two pairs of cysteines (or histidines) are thought to form a loop or 'finger' of protein that contacts DNA. This hypothesis was initially suggested by the observation that the RNA polymerase III transcription factor A (TFIIIA) of *Xenopus laevis*², which binds to both DNA and RNA, contains nine zinc binding finger motifs and approximately nine bound zinc ions³.

The wide variety of DNA binding proteins that contain the proposed zinc binding finger include several transcriptional regulatory proteins from the yeast *S. cerevisiae* and one from

α Lys Cys Ser
 Leu Lys
 Lys Glu
 Lys Lys
 Leu **Pro** → Leu
 Arg Lys
 Cys Cys
 Ile Ala
 Asp Lys
 NH₂—Ala Cys Leu—COOH
 Zn

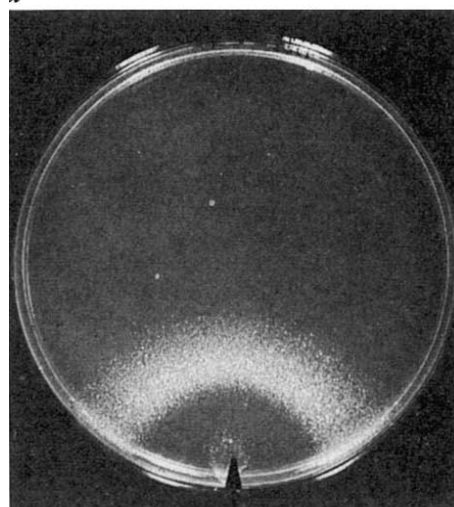
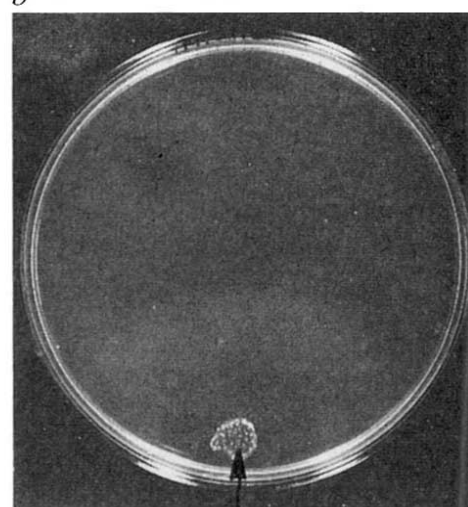
b

GAL4	10	Ala	Cys	Asp	Ile	Cys	Arg	Leu	Lys	Lys	Leu	Lys	Cys	Ser	Lys	Glu	Lys	Pro	Lys	Cys	Ala	Lys	Cys	Leu
LAC9	94	Ala	Cys	Asp	Ala	Cys	Arg	Lys	Lys	Lys	Trp	Lys	Cys	Ser	Lys	Thr	Val	Pro	Thr	Cys	Thr	Asn	Cys	Leu
PPR1	33	Ala	Cys	Lys	Arg	Cys	Arg	Leu	Lys	Lys	Ile	Lys	Cys	Asp	Gln	Glu	Phe	Pro	Ser	Cys	Lys	Arg	Cys	Ala
ARGR2	20	Gly	Cys	Trp	Thr	Cys	Arg	Gly	Arg	Lys	Val	Lys	Cys	Asp	Leu	Arg	His	Pro	His	Cys	Gln	Arg	Cys	Glu
Qa-1F	75	Ala	Cys	Asp	Gln	Cys	Arg	Ala	Ala	Arg	Glu	Lys	Cys	Asp	Gly	Ile	Gln	Pro	Ala	Cys	Phe	Pro	Cys	Val
Consensus		Ala	CYS	Asp	x	CYS	ARG	Pho	Lys	Lys	Arg	Arg	Pho	LYS	CYS	Asp	Ser	x	x	x	PRO	x	CYS	Pho
		4	5	3		5	5	4	4	5	3	5	5	5		5	5			5	5	3	5	4

One mutation that inactivates the GAL4-encoded protein (Pro 26 $\rightarrow$ Leu) alters the proline residue located two amino acids before the second pair of cysteines that are thought to be responsible for zinc binding (see Fig. 1a). This proline residue is conserved in all five proteins whose sequences are shown in Fig. 1b. A yeast strain carrying this *gal4* mutation is Gal⁻, unable to grow on media containing galactose as the sole carbon source.

The ability of zinc to correct the DNA binding defect of the Pro 26→Leu mutant protein was also measured *in vitro*. The mutant protein was expressed in *Escherichia coli* and the DNA binding activity of the GAL4 protein in extracts of these cells

2


$$\text{ZnCl}_2$$
$$b$$
 ZnCl_2

acetate. Growth of this mutant is not stimulated by CaCl_2 , MgCl_2 , NiCl_2 , CdCl_2 , CoCl_2 , FeSO_4 , CuSO_4 , MnSO_4 or NaCl . *b*, Inability of ZnCl_2 to stimulate growth of the *Leu 32* $\rightarrow$ *Pro gal4* mutant. Strain YM1188 (CTG $\rightarrow$ CCG) was treated as described above. Similar results were obtained with 13 other *gal4* mutations⁹ that alter amino acids in and adjacent to the zinc binding finger.

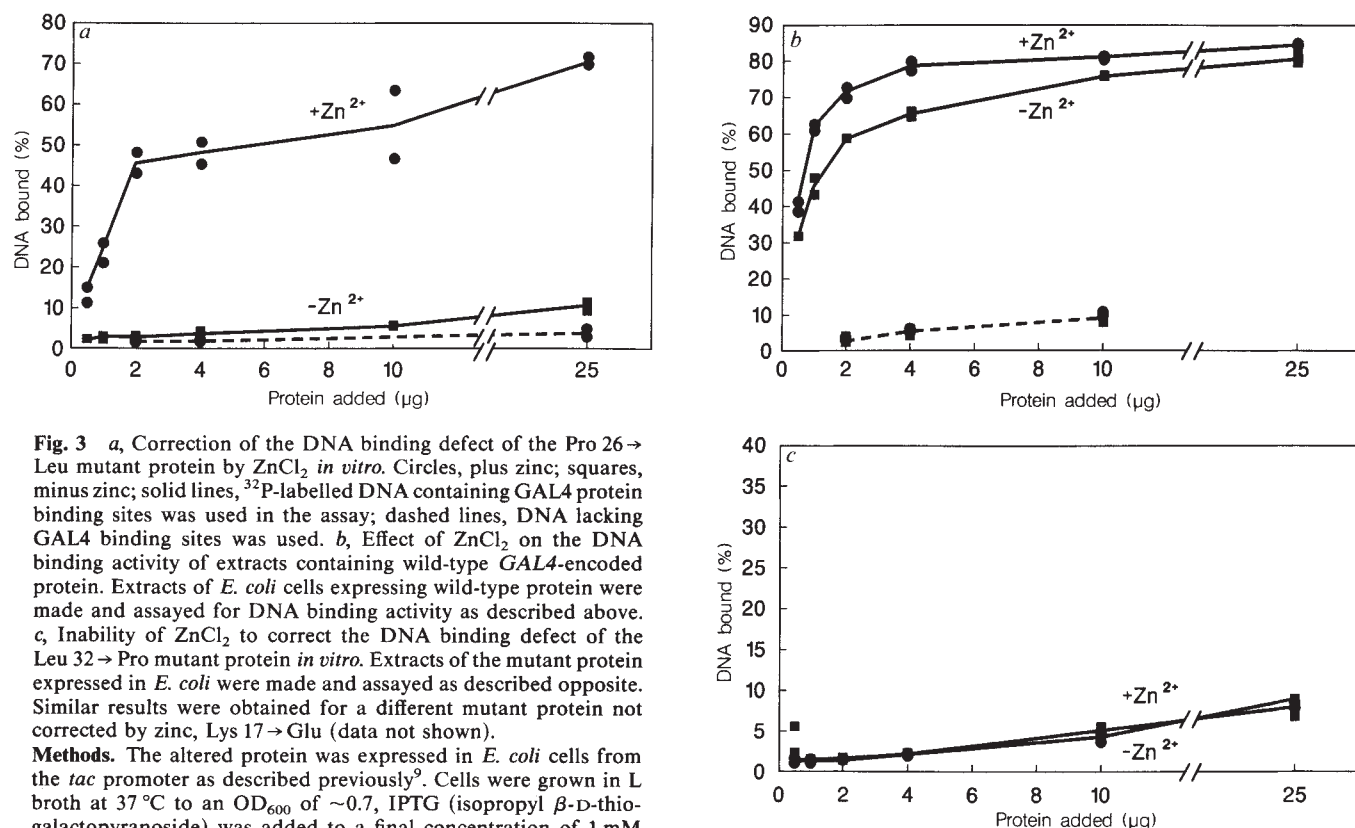


Fig. 3 *a*, Correction of the DNA binding defect of the Pro 26 → Leu mutant protein by ZnCl₂ *in vitro*. Circles, plus zinc; squares, minus zinc; solid lines, ³²P-labelled DNA containing GAL4 protein binding sites was used in the assay; dashed lines, DNA lacking GAL4 binding sites was used. *b*, Effect of ZnCl₂ on the DNA binding activity of extracts containing wild-type GAL4-encoded protein. Extracts of *E. coli* cells expressing wild-type protein were made and assayed for DNA binding activity as described above. *c*, Inability of ZnCl₂ to correct the DNA binding defect of the Leu 32 → Pro mutant protein *in vitro*. Extracts of the mutant protein expressed in *E. coli* were made and assayed as described opposite. Similar results were obtained for a different mutant protein not corrected by zinc, Lys 17 → Glu (data not shown).

Methods. The altered protein was expressed in *E. coli* cells from the *tac* promoter as described previously⁹. Cells were grown in L broth at 37 °C to an OD₆₀₀ of ~0.7, IPTG (isopropyl β-D-thiogalactopyranoside) was added to a final concentration of 1 mM (to induce expression from the *tac* promoter) either with or without 0.1 mM (final concentration) ZnCl₂. Two hours later cells were harvested and lysed by sonication (10 bursts of 2 s, each followed by 10 s of cooling) in 1/100 volume of buffer D[200] either with or without 2 mM ZnCl₂. (Buffer D is 25 mM Tris pH 7.5, 5 mM MgCl₂, 1 mM EDTA, 12 mM 2-mercaptoethanol, 10% (w/v) glycerol. The number in square brackets is the KCl concentration in mM.) Cleared lysates, prepared by centrifugation in a Sorval SS34 rotor for 3 h at 15,000 r.p.m., were frozen in liquid N₂ and stored at -80 °C. Protein concentration of the extracts (usually ~20 mg ml⁻¹) was determined by the method of Bradford¹². The DNA binding activity of the extracts was determined by a nitrocellulose filter binding assay¹⁰ using buffer D[50] either in the presence or absence of 2 mM ZnCl₂, essentially as described previously⁹. Values are the proportion of ³²P-labelled DNA that is bound to the filters (%) as a function of extract protein added to the assay. Solid lines, pBM992 (ref. 9), consisting of pUC18 + the GAL1-10 control region, which contains 4 binding sites for the GAL4-encoded protein^{13,14}; dotted lines, pUC18.

was assayed by a nitrocellulose filter binding assay¹⁰, as described previously⁹. Figure 3*a* shows that GAL4 protein from the Pro 26 → Leu mutant has substantial DNA binding activity when extracts of *E. coli* cells, induced for GAL4 expression in the presence of ZnCl₂, are made in buffer containing ZnCl₂ and assayed in the presence of ZnCl₂; extracts made and assayed without added ZnCl₂ have negligible DNA binding activity. The DNA binding activity of wild-type protein is unaffected by the presence of ZnCl₂ (Fig. 3*b*); ZnCl₂ has no effect on the *in vitro* DNA binding activity of GAL4 protein from a mutant whose growth defect is not corrected by zinc (Fig. 3*c*). These results strengthen the conclusion that the zinc binding finger of GAL4-encoded protein indeed contains zinc ions that are essential for its DNA binding activity. They also imply that sequences similar to the zinc binding finger of GAL4 protein in other DNA binding and transcriptional regulatory proteins are indeed zinc-containing DNA-binding domains.

Our results suggest that a GAL4 protein containing leucine instead of proline at position 26 binds zinc less readily than wild-type protein. It is not unreasonable to imagine that this proline residue causes a bend in the protein that brings the two pairs of cysteine residues close enough together to chelate a zinc ion. The protein that does not have this bend (Pro 26 → Leu) might have a lower affinity for zinc because close apposition of the pairs of cysteines is less favourable. We expect to find other *gal4* mutations that are corrected by high concentrations of zinc;

their analysis is likely to lend insight into the structure and function of the zinc binding finger. The analysis of zinc-remedial mutants provides a powerful tool for elucidating the structure and function of the zinc binding finger that complements information obtained from sequence homology searches. Furthermore, this genetic approach could be used to identify genes encoding other zinc-containing proteins whose sequences are presently unknown.

I thank my colleagues Athanasios Theologis, David Schlessinger, Bob Waterston, Don Moerman, John Majors, Maynard Olson, Peter Burgers and Jim Dover for helpful suggestions. This work was supported by a grant from the N.I.H. M.J. is an Established Investigator of the American Heart Association.

Received 2 April; accepted 29 May 1987.

1. Berg, J. *Science* **232**, 485-487 (1986).
2. Ginsberg, A. M., King, B. O. & Roeder, R. G. *Cell* **39**, 479-489 (1984).
3. Miller, J., McLachlan, A. D. & Klug, A. *EMBO J.* **4**, 1609-1614 (1985).
4. Laughon, A. & Gesteland, R. F. *Molec. cell. Biol.* **4**, 260-267 (1984).
5. Salmeron, J. M. Jr & Johnston, S. A. *Nucleic Acids Res.* **14**, 7767-7781 (1986).
6. Kammerer, B., Guyonvarch, A. & Hubert, J. C. *J. molec. Biol.* **180**, 239-250 (1984).
7. Messenguy, R., Dubois, E. & Descamps, F. *Eur. J. Biochem.* **157**, 77-81 (1986).
8. Baum, J. A., Geever, R. & Giles, N. H. *Molec. cell. Biol.* **7**, 1256-1266 (1987).
9. Johnston, M. & Dover, J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2401 (1987).
10. Riggs, A. D., Suzuki, H. & Bourgeois, S. *J. molec. Biol.* **48**, 67-83 (1970).
11. Johnston, M. & Davis, R. W. *Molec. cell. Biol.* **4**, 1440-1448 (1984).
12. Bradford, M. M. *Analyt. Biochem.* **72**, 248-254 (1976).
13. Bram, R. & Kornberg, R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 43-47 (1985).
14. Giniger, E., Varnum, S. & Ptashne, M. *Cell* **40**, 767-774 (1985).

A distant enhancer element is required for polymerase III transcription of a U6 RNA gene

Christina Bark, Polly Weller, Jan Zabielski, Lars Janson & Ulf Pettersson*

Department of Medical Genetics, Biomedical Center, Box 589, S-57123 Uppsala, Sweden

U RNAs are highly abundant small nuclear RNAs involved in the processing of messenger RNA^{1,2}. Most U RNA genes are thought to be transcribed by RNA polymerase II (pol II). However, evidence has recently been presented that U6 RNA genes are transcribed by RNA polymerase III (pol III)^{3,4}. In the light of these results it was surprising to find that the 5' flanking region of a mouse U6 RNA gene⁵ includes a perfect copy of the octamer sequence motif, ATTTGCAT, found in many RNA polymerase II transcription enhancer elements⁶⁻⁸. In the present study we show that deletion of mouse U6 gene sequences upstream of nucleotide position -217, including the octanucleotide motif, reduces U6 transcription by 90% when assayed in *Xenopus laevis* oocytes, suggesting the presence of a distant control element. DNase I footprinting of the 5' flanking region of the U6 gene shows protection of the octanucleotide sequence. Moreover, the 5' flanking sequence from -217 to -315 can replace the enhancer of a human U2 RNA gene. We therefore conclude that although U6 RNA genes appear to be transcribed by pol III, they are preceded by an enhancer-like element which can functionally substitute for the enhancer of a pol II-transcribed U RNA gene.

Previous studies of U2 and U4 RNA genes have shown that their enhancers contain a single copy of a conserved octameric sequence, ATTTGCAT, which may occur in either orientation⁷⁻¹¹, as well as other distinct sequence motifs involved in DNA-protein interactions (L.J., P.W., C.B. and U.P., unpublished results). Sequence comparison between a mouse U6 RNA gene and human genes for U2 and U4 RNAs revealed two 5' flanking sequence similarities: one is in the -50 to -60 region¹¹, and the other is upstream of -223 and includes one perfect copy of the octamer motif, ATTTGCAT (Fig. 1a). The octamer motif, first described in the transcriptional control elements of immunoglobulin genes^{6,12}, has also been found in the 5' flanking sequences of many other genes⁷⁻¹¹ transcribed by pol II and has been demonstrated specifically to bind a distinct protein factor^{7,8,13-15}. To investigate the possibility that the mouse U6 gene contains an upstream regulatory element even though transcribed by pol III, we carried out transcription analysis of various promoter constructs (Fig. 1b) by injection into *Xenopus laevis* oocytes (Fig. 1c and d). This assay system has previously been used for the transcription analysis both of U6 (ref. 4) and other U RNA genes^{9-11,16-18}. To eliminate any possible influence by uncharacterized sequences, a fragment of plasmid pMU6.52BE (ref. 5) of ~600 base pairs (bp) was subcloned into the pGEM2 vector to give plasmid pMU6.wt. This plasmid contains the U6 coding sequence and 315 bp of 5' flanking sequence (Fig. 1b). Oocyte injections showed that U6 RNA is transcribed with equal efficiency from templates pMU6.52BE and pMU6.wt in *X. laevis* oocytes (data not shown). Transcription of the two U6 templates in the presence of α -amanitin confirmed that U6 RNA synthesis is dependent on a pol III-like activity (Fig. 1c). In the construct pMU6.dSspI all U6 sequences upstream from the SspI site at -217, including the octamer motif, were deleted. Transcription from this template was reduced by ~90% compared to pMU6.wt (Fig. 1c). The sequence deleted from pMU6.dSspI must therefore contain an

element required for efficient transcription of the U6 RNA gene.

To determine if this deleted region interacted with DNA-binding proteins, we used an electrophoretic mobility shift assay^{13,19-21}. An end-labelled fragment, extending from -217 to -315 of the mouse U6 5' flanking sequence, was assayed for ability to bind proteins using a HeLa cell nuclear extract²². Distinct DNA-protein complexes were seen following electrophoresis in native polyacrylamide gels (Fig. 2a). Different unlabelled DNA fragments were added in excess to demonstrate that the binding was specific. The two most prominent complexes (A and B in Fig. 2a) could be competed out using an 18-bp double-stranded oligonucleotide homologous with a sequence in the human U4C gene enhancer and which contains one copy of the octamer motif (Fig. 2b). Similar competition of A and B was also observed with a 75-bp fragment from the mouse κ light-chain promoter which also contains the octamer motif. Vector DNA fragments (pGEM2) lacking homology to the octamer element were unable to compete (data not shown).

To demonstrate further that the association between protein factors and mouse U6 upstream sequences included a specific interaction with the octamer motif, we carried out DNase I footprint analysis³⁰ with a HeLa cell nuclear extract²³, using a 226-bp DNA fragment extending from the *Dra*I site at position -89 to the *Bam*HI site at -315. There was a distinct footprint covering at least 18 nucleotides (nt) on both strands, spanning the region of the octamer motif (Fig. 3a). The protected region is shown schematically in Fig. 3c. To confirm that the observed protection was caused by specific binding of the octamer factor, footprint reactions were carried out in the presence of an excess of different competitor DNA fragments. Two different fragments containing the octamer motif were used; one derived from the mouse immunoglobulin κ light-chain gene and the other from a human U4C gene¹¹ (Fig. 2b). The octamer-binding factor binds to the consensus motif in both sequence contexts (ref. 13; P.W. and C.B., unpublished results). Protection from DNase I cleavage was abolished in the presence of these competitor DNA fragments. In contrast, the footprint was unaltered when a nonspecific DNA fragment (pBR325) was used as a competitor (Fig. 3b). We therefore conclude that the octamer factor binds to the motif in the 5' flanking sequence of the mouse U6 RNA gene.

To demonstrate that the U6 upstream element does indeed have an enhancer-like function, plasmid pHU2.MU6 was constructed, in which the human U2 gene enhancer was replaced by the mouse U6 gene sequence between -316 and -217 (Fig. 1b). The transcriptional activity of this recombinant plasmid was assayed in *X. laevis* oocytes and compared with those of two human U2 gene plasmid constructs; pBH1, which contains the U2 enhancer and is transcribed at the wild-type level, and pBS4, in which the entire enhancer has been deleted and which is transcribed at a low level (~10% as much as pBH1) in *X. laevis* oocytes¹⁸ (Fig. 1b and d). The mouse U6 upstream element enhanced the synthesis of U2 RNA (predominantly, but not exclusively, the precursor form) to ~50-60% of the level obtained with pBH1, thus demonstrating the presence of an enhancer-like element (Fig. 1d). Furthermore, this shows that the mouse U6 upstream element can substitute for the transcriptional enhancer of a human U2 RNA gene even though the latter is transcribed by pol II.

Results from a large number of studies have led to the conclusion that the three known mammalian RNA polymerases have different promoter requirements. No indications for transcription factors common to different polymerases have been presented to date. Pol III seems to differ from pol I and pol II in being dependent primarily on intragenic control elements, as has been clearly demonstrated in studies of transcription of 5S RNA genes^{24,25}. However, promoters in other pol III genes, for example VA RNA genes²⁶ from adenovirus and transfer RNA genes²⁷, include regulatory sequences which extend into the immediate 5' flanking region. Sequences close to the initiation

* To whom correspondence should be addressed.

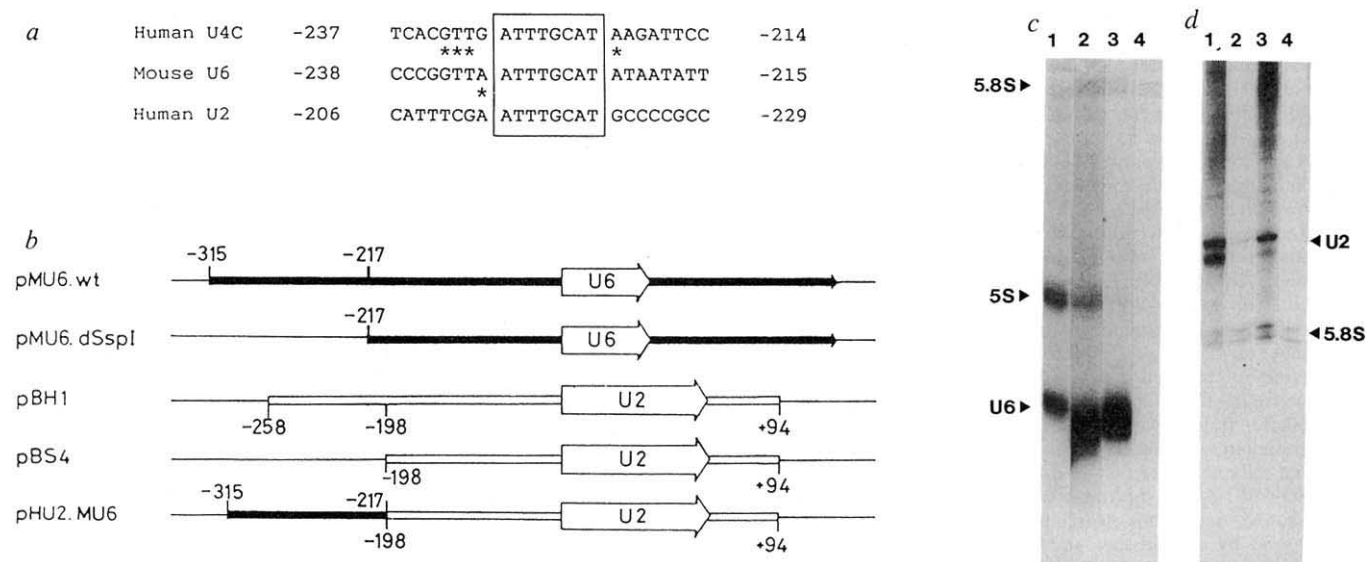


Fig. 1 Transcription analysis of U RNA gene recombinant plasmids in *Xenopus laevis* oocytes. **a**, Sequence of parts of the enhancer regions of human U2, U4C and mouse U6 genes with the octamer motif aligned (boxed). The motif is in the opposite orientation in the U2 gene. Asterisks, additional similarities in the sequence immediately flanking the octamer motif. **b**, Schematic representation of plasmid constructs used in transcription analysis. Open arrows, U6 and U2 coding sequences; solid and open bars, sequences from the mouse U6 and human U2 gene loci, respectively; single line, vector sequences. A 600-bp *Bam*HI-*Hind*III fragment from clone pMU6.52BE (ref. 5) was subcloned into the *Bam*HI and *Hind*III sites of the pGEM2 vector (Promega-Biotech) using standard procedures. This clone, designated pMU6.wt, contains the U6 gene with 315 bp of 5'-flanking sequence and ~200 bp of 3' flanking sequence. Similarly, a *Ssp*I-*Hind*III fragment, including upstream sequences to -217 was subcloned into the *Sma*I and *Hind*III sites of the same vector to give pMU6.dSspI. Plasmid pBH1 contains the coding region for human U2 RNA and 258 pb of 5' flanking sequence, which includes the U2 enhancer. In pBS4 the 5' flanking sequence extends only to position -198, deleting the enhancer element. The construction of pBH1 and pBS4 has been described previously¹⁸. Recombinant pHU2.MU6 was constructed by ligating the *Sma*I-*Hind*III fragment from pBH1, which spans from -198 to +94 in the U2 gene, together with a U6 *Bam*HI-*Ssp*I fragment (-217 to -315) from pMU6.wt, into the pGEM2 vector cleaved with *Bam*HI and *Hind*III. Recombinant constructs were verified by DNA sequencing. **c**, Transcriptional analysis of U6 recombinant plasmids. Oocytes were microinjected with the following RNA templates: lane 1, pMU6.dSspI; lane 2, pMU6.wt; lane 3, pMU6 with 1 μ g ml⁻¹ α -amanitin; lane 4, pMU6.wt with 200 μ g ml⁻¹ α -amanitin. The migration positions of U6 and the endogenous 5S and 5.8S RNAs are indicated. The identity of the transcript arising from this U6 mouse template has been verified by fingerprint analysis⁴. The transcripts from pMU6.dSspI are a few nucleotides longer than those from pMU6.wt. This may represent incomplete 3' processing, as with U2 5' flanking sequence deletion mutants¹⁸. In the presence of 1 μ g ml⁻¹ α -amanitin the U6 template DNA represses endogenous 5S RNA transcription, suggesting competition for limiting pol III transcription factors. **d**, Transcriptional analysis of U2 RNA recombinant plasmids. The migration positions of U2 and the endogenous 5.8S RNA are indicated. Oocytes were injected with the following templates: lane 1, pBH1; two transcripts are seen which correspond to the mature U2 RNA (189 nt) and a precursor (200 nt); lane 2, pBS4; only the precursor form is present; lane 3, pHU2.MU6: both precursor and mature forms are present; lane 4, pHU2.MU6 with 1 μ g ml⁻¹ α -amanitin.

Methods. *X. laevis* oocytes were injected with 1 ng of supercoiled plasmid templates as has been described^{16,17}. Where appropriate, α -amanitin was co-injected with plasmid DNA to give final concentrations of either 1 μ g ml⁻¹ or 200 μ g ml⁻¹ in the nucleus. RNA was extracted from individual oocytes after 20 h incubation at 20 °C and analysed by electrophoresis in denaturing 10% polyacrylamide gels (20:1). Gels were autoradiographed at -70 °C with intensifying screens. Scanning densitometry was used to measure the relative transcript levels.

Fig. 2 Binding of a nuclear factor to mouse U6 gene upstream sequences. **a**, An *Ssp*I-*Bam*HI fragment extending from -217 to -315 of the mouse U6 5' flanking sequence was incubated in HeLa cell nuclear extract²² as described below. The bound protein-DNA complexes and free DNA were resolved by native polyacrylamide gel electrophoresis. Lane 1, free DNA control; lane 2, no competitor; lanes 3 and 4, 20 and 40 times molar excess, respectively, of the 18-bp U4C gene oligonucleotide (**b**); lanes 5 and 6, 100 and 300 times molar excess, respectively, of a 75-bp *Eco*RI-*Pvu*II fragment from the mouse immunoglobulin κ light chain MOPC-41 gene¹³ (gift of H. Singh), which includes the octamer motif. The strong band at the bottom of each lane is the unbound DNA fragment. Solid triangles, major retarded complexes (A and B); open triangle, a new complex arising when the MOPC-41 gene fragment is used as a competitor, the identity of which is not known. Additional complexes are not the result of binding of the octamer factor. **b**, Sequence (upper strand) of the 18-bp double-stranded oligonucleotide used as a competitor in the electrophoretic mobility shift assay and DNase I footprint analysis. The oligonucleotide is identical in sequence with the human U4C gene enhancer region between -217 and -234 (ref. 11), and contains one copy of the octamer motif which is shown in the box.

Methods. Binding analysis: the *Ssp*I-*Bam*HI fragment was end-labelled using [α -³²P]dCTP (3,000 Ci mmol⁻¹) and the Klenow fragment of *Escherichia coli* DNA polymerase I and was then purified by electrophoresis in low-melting agarose. Binding reactions (10 μ l) contained 0.3 ng fragment (5,000 c.p.m.), 10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 1 mM dithiothreitol (DTT), 1 mM EDTA, 5% glycerol and 3 μ g of HeLa cell nuclear protein extract²². Reactions 2-6 also contained 2 μ g of poly(dI-dC).poly(dI-dC) and the appropriate competitor DNA. After incubation at 30 °C for 30 min, the reactions were electrophoresed directly in a 4% polyacrylamide gel (30:1) containing 50 mM Tris-HCl, 380 mM glycine, 2 mM EDTA (pH 8.5). Gels were dried and autoradiographed with intensifying screens.

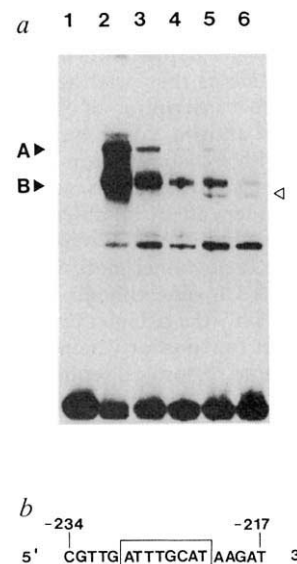
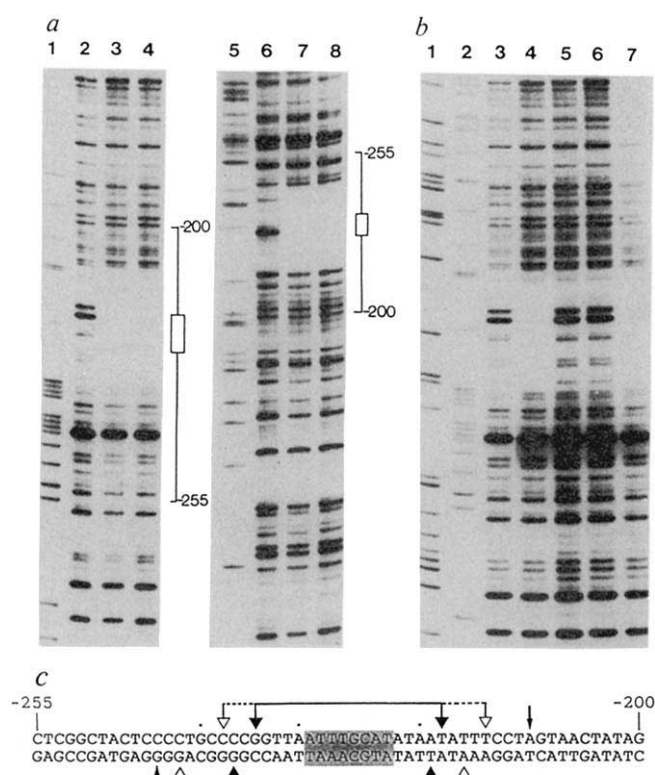


Fig. 3 DNase I protection footprinting of mouse U6 5' flanking sequences in HeLa cell nuclear extract. *a*, Non-coding (lanes 1–4) and coding strands (lanes 5–8) kinase labelled at the *Bam*HI (–315) and *Dra*I (–89) sites, respectively. Lanes 1 and 5, C reaction sequence ladder of the probe fragments; lanes 2 and 6, no protein extract; lanes 3 and 7, and 4 and 8, reactions contained 10 µg extract digest with 200 and 250 ng DNase I, respectively. Positions –200 and –255 in the 5' flanking sequence, together with the octameric consensus sequence shown as an open box, are indicated to the right of each panel. No differences in protection were seen using larger amounts of nuclear extract (data not shown). *b*, Footprint competition (non-coding strand): lanes 1 and 2, G and C sequence ladders of the probe fragment; lane 3, no extract. Reactions shown in lanes 4–7 contained 10 µg extract and the following competitor DNA fragments: lane 4, no competitor; lane 5, 150 times molar excess of the human U4C gene octamer oligonucleotide (Fig. 2*b*); lane 6, 150 times molar excess of a mouse immunoglobulin MOPC-41 κ light chain 75-bp *Pvu*I–*Eco*RI promoter fragment; lane 7, 50 ng linearized pBR325. *c*, Schematic representation of U6 sequences protected from DNase I digestion in HeLa cell nuclear extract. The 5' flanking sequence between positions –200 and –255 is shown on both strands. The octanucleotide motif is shaded. Sequences clearly protected from DNase I cleavage are enclosed by a solid line and solid arrows. Open arrows, the first unprotected base on either side of the footprint (indicates the maximum size of the footprint); small arrows, sites of enhanced DNase I cleavage.

Methods. Preparation of end-labelled DNA fragments: pMU6.52BE was cleaved with *Bam*HI or *Dra*I and phosphorylated with [γ - 32 P]dATP (3,000 Ci mmol $^{-1}$) and polynucleotide kinase. DNA was recleaved with *Dra*I or *Bam*HI, respectively, and fragments purified following electrophoresis in low melting-point agarose. HeLa cell nuclear extract was prepared from Spinner cultures as described²³, except that a single protein precipitation step was carried out by the addition of (NH $_4$) $_2$ SO $_4$ to a final concentration of 0.25 g ml $^{-1}$. DNase I footprinting was carried out essentially as described²³. The standard reaction contained the following components in a final volume of 50 µl: 25 mM HEPES (pH 7.8), 50 mM KCl, 0.05 mM EDTA, 0.5 mM DTT and 5% glycerol. Reaction mixes contained 20 ng µl $^{-1}$ poly(dI-dC).poly(dI-dC), 20,000 c.p.m. end-labelled fragment (1–4 ng) and 10 µg protein extract (except control reactions). Preincubation in the presence of poly(dI-dC).poly(dI-dC) and competitor DNA fragments, where appropriate, was at 25 °C for 10 min. The end-labelled fragment was added and incubation continued for 15 min. MgCl $_2$ was added to 5 mM, followed by DNase I (200 ng in the presence of extract and 80 ng in the absence of extract, unless otherwise stated). Digestion was allowed to proceed for 60 s at room temperature and was terminated by the addition of 100 µl of a stop solution (100 mM Tris-HCl, pH 7.9, 100 mM NaCl, 10 mM EDTA, 1% sarkosyl, 500 µg ml $^{-1}$ proteinase K and 30 µg ml $^{-1}$ salmon sperm DNA. Reactions were incubated for 30 min at 65 °C before phenol extraction and ethanol precipitation. DNA fragments were analysed in denaturing 6% polyacrylamide gels (29:1 crosslink).



site have also been shown to have an effect on the transcription of a 7SL RNA gene²⁸. Significantly, in no cases have distant sequences with enhancer-like function been implicated in pol III transcriptional control. Indeed, the polyoma virus enhancer has been shown to have no effect on transcription when introduced upstream from the adenovirus *VAI* gene²⁹.

Several criteria, in addition to sensitivity to α -amanitin, have been used to show that U6 genes are transcribed by pol III. A human U6 gene has been shown to be transcribed in an S100 extract depleted of pol II activity³ and a mouse U6 gene (the same as that used here) can compete for factors required for the transcription of 5S RNA by pol III (ref. 4). In addition, the La antigen, which binds to most pol III products, is associated with transcripts from both U6 genes^{3,4}. This study shows that transcription of a mouse U6 gene by an RNA pol III-like activity is dependent on an enhancer-like element located upstream of position –217. We have shown that one or more factors bind to the octamer motif in this region. Site-specific mutagenesis of the U6 gene enhancer will be necessary to discover whether it is only the octamer binding protein that is important for enhancer function or whether other DNA–protein interactions in the same region are required for transcription.

The U6 enhancer element can functionally substitute for the enhancer of a gene transcribed by pol II, suggesting that pol II and pol III can use common transcription factors. The distinction between transcription by pol II and pol III is therefore less clear than hitherto believed. It remains possible, however, that U6 RNA genes are transcribed by a new or modified RNA polymerase. U6 differs from other pol III products in having a

monomethylated (and as yet unidentified) cap structure, suggesting that the transcription mechanism may be different. However, if this is the case, it is likely to be closely related to pol III.

We thank Dr Hans Hultberg (Kabigen AB) for a synthetic oligodeoxynucleotide, Dr Y. Ohshima for the mouse U6 gene, Dr H. Singh for a plasmid containing the mouse κ light-chain promoter, all members of Dr P. A. Sharp's laboratory for technical advice, Drs G. Akusjärvi and M. Ustav for critical reading of the manuscript, Britt-Marie Carlberg for technical assistance and Katarina Eriksson for secretarial aid. This investigation was supported by grants from the Swedish Medical Research Council, the Swedish National Board for Technical Development and the Marcus Borgström Foundation grants from Uppsala University. P.W. was supported by a fellowship in the Royal Society European Science Exchange Programme. L.J. is a fellow of the Swedish Cancer Society.

Note added in proof: Since submission of this paper, Krol *et al.*³¹ have observed the presence of an octamer sequence motif in the 5' flanking sequence of a *Xenopus tropicalis* U5 RNA gene.

Received 25 March; accepted 13 May 1987.

- Maniatis, T. & Reed, R. *Nature* **325**, 673–678 (1987).
- Sharp, P. A. *Science* **235**, 766–771 (1987).
- Kunkel, G. R., Maser, R. L., Calvet, J. P. & Pederson, T. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8575–8579 (1986).
- Reddy, R., Henning, D. & Das, G. *J. biol. Chem.* **262**, 75–81 (1987).
- Ohshima, Y., Okada, N., Tani, T., Itoh, Y. & Itoh, M. *Nucleic Acids Res.* **9**, 5145–5158 (1981).
- Falkner, F. G. & Zachau, H. G. *Nature* **310**, 71–74 (1984).
- Sive, H. L. & Roeder, R. G. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6382–6386 (1986).
- Bohmann, D. *et al.* *Nature* **325**, 268–272 (1987).
- Mattaj, J. W., Leinhardt, S., Jiricny, J. & De Robertis, E. M. *Nature* **316**, 163–167 (1985).

10. Ares, M., Mangin, M. & Weiner, A. M. *Molec. cell. Biol.* **5**, 1560-1570 (1985).
11. Bark, C., Weller, P., Zabielski, J. & Pettersson, U. *Gene* **50**, 333-344 (1986).
12. Parslow, T. G., Blair, D. L., Murphy, W. J. & Granner, D. K. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2650-2654 (1984).
13. Singh, H., Sen, R., Baltimore, D. & Sharp, P. A. *Nature* **319**, 154-158 (1986).
14. Landolfi, N. F., Capra, J. D. & Tucker, P. W. *Nature* **323**, 548-551 (1986).
15. Staudt, L. M. *et al. Nature* **323**, 640-643 (1986).
16. Murphy, J. T., Burgess, R. R., Dahlberg, J. E. & Lund, E. *Cell* **29**, 265-274 (1982).
17. Skuzeski, J. M. *et al. J. biol. Chem.* **259**, 8345-8352 (1984).
18. Westin, G., Lund, E., Murphy, J. T., Pettersson, U. & Dahlberg, J. E. *EMBO J.* **3**, 3295-3301 (1984).
19. Fried, M. & Crothers, D. M. *Nucleic Acids Res.* **9**, 6505-6525 (1981).
20. Garner, M. M. & Revzin, A. *Nucleic Acids Res.* **9**, 3047-3060 (1981).
21. Carthew, R. W., Chodosh, L. A. & Sharp, P. A. *Cell* **43**, 439-448 (1984).
22. Dignam, J. D., Lebowitz, R. M. & Roeder, R. G. *Nucleic Acids Res.* **11**, 1475-1489 (1983).
23. Ohlsson, H. & Edlund, T. *Cell* **45**, 35-44 (1986).
24. Sakonju, S., Bogenhagen, D. F. & Brown, D. D. *Cell* **19**, 13-25 (1980).
25. Bogenhagen, D. F., Sakonju, S. & Brown, D. D. *Cell* **19**, 27-35 (1980).
26. Thimmappaya, B., Jones, N. & Shenk, T. *Cell* **18**, 947-954 (1979).
27. Sharp, S. J., Schaack, J., Cooley, L., Burke, D. J. & Söll, D. *CRC crit. Rev. Biochem.* **19**, 107-144 (1985).
28. Ullu, E. & Weiner, A. M. *Nature* **318**, 371-374 (1985).
29. Berger, S. L. & Folk, W. R. *Nucleic Acids Res.* **13**, 1413-1428 (1985).
30. Galas, D. & Schmitz, A. *Nucleic Acids Res.* **5**, 3157-3170 (1978).
31. Krol, A., Carbon, P., Ebel, J.-P. & Appel, B. *Nucleic Acids Res.* **15**, 2463-2478 (1987).

Modulation of spectrin-actin assembly by erythrocyte adducin

Kevin Gardner & Vann Bennett

Department of Cell Biology and Anatomy, Johns Hopkins School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, USA

The spectrin-based membrane skeleton, an assembly of proteins tightly associated with the plasma membrane, determines the shape and mechanical properties of erythrocytes. Spectrin, the most abundant component of this assembly, is an elongated and flexible molecule that, with potentiation by protein 4.1, is cross-linked at its ends by short actin filaments to form a lattice beneath the membrane. These and other proteins stabilize the plasma membrane, organize integral membrane proteins and maintain specialized regions of the cell surface¹⁻³. A membrane-skeleton-associated calmodulin-binding protein of erythrocytes⁴ is a major substrate for Ca²⁺- and phospholipid-dependent protein kinase C (ref. 5), and thus is a target for Ca²⁺ by two regulatory pathways. Here we demonstrate that this protein, called adducin: (1) binds tightly *in vitro* to spectrin-actin complexes but with much less affinity either to spectrin or to actin alone; (2) promotes assembly of additional spectrin molecules onto actin filaments; and (3) is inhibited in its ability to induce the binding of additional spectrin molecules to actin by micromolar concentrations of calmodulin and Ca²⁺. Adducin may be involved in the action of Ca²⁺ on erythrocyte membrane skeleton and in the assembly of spectrin-actin complexes.

We evaluated the ability of adducin to associate with either spectrin, rabbit skeletal muscle actin or a mixture of spectrin and actin by rate-zonal sedimentation of these proteins on sucrose gradients (Fig. 1A). In the presence of spectrin and actin, adducin sediments to the bottom of the centrifuge tube together with actin filaments (F-actin), whereas adducin sediments independently in the presence of either actin or spectrin. Thus, adducin associates with spectrin-actin complexes but shows much lower affinity for either protein alone. Measurements of adducin binding to spectrin immobilized on beads (not shown) indicate a weak association between these proteins that is at least tenfold lower in affinity than binding of adducin to spectrin-actin complexes (see below).

A spectrin-dependent interaction between adducin and actin for the formation of a stable adducin-spectrin-actin oligomer is supported by further experiments. First, we manipulated the concentration of spectrin-actin complexes without varying the concentration of spectrin or actin by comparing spectrin from humans (low affinity for actin) with spectrin from pigs or sheep (higher affinity for actin)^{6,7}. Despite the species differences, pig spectrin is about fivefold more active than human spectrin in

promoting binding of human adducin to actin (Fig. 1B); we obtained equivalent results with sheep spectrin (not shown). In the subsequent studies we used pig spectrin because of enhanced binding of adducin compared with human spectrin. In a second type of experiment we evaluated binding of adducin to actin in the presence and absence of pig spectrin (Fig. 1C). Binding of adducin to actin alone is linear with increasing concentrations of adducin, as expected for a low-affinity, high-capacity binding interaction. We estimate a minimum K_d for adducin binding to actin alone as 28 μ M. Such behaviour would also be observed for nonspecific binding, and further studies are required to verify the biological significance of binding of adducin to actin in isolation. In contrast, spectrin-dependent binding of adducin to actin (Fig. 1D) is saturable and half-maximal in 80 nM adducin. The saturable binding of adducin is not caused by limitations in the amount of spectrin, as at least 60% of spectrin is not associated with actin, nor is it caused by limitations in the binding sites on actin filaments as >90% of the potential spectrin-binding sites on actin are unoccupied.

Adducin also increases binding of spectrin to actin. We measured the effect of adducin on spectrin association with actin using a fixed concentration of ¹²⁵I-labelled spectrin in the presence of various concentrations of unlabelled adducin (Fig. 2). Half-maximal stimulation of spectrin binding to actin occurs with 150 nM adducin, and maximal binding is limited to about twice the level of spectrin bound in the absence of adducin, even though nearly 80% of the total spectrin remains free. It is unlikely that adducin promotes binding of spectrin by an effect on spectrin self-association, because adducin has no influence on the spectrin dimer-tetramer equilibrium when analysed by native polyacrylamide gel electrophoresis (PAGE) (not shown).

Adducin-dependent binding of spectrin to actin could result from two mechanisms: (1) adducin stabilizes spectrin-actin complexes; or (2) a new site for spectrin is created by association of adducin with spectrin and actin, with the consequence that two (or more) spectrin molecules are involved with each adducin on actin filaments. Some circumstantial evidence in favour of a new site for spectrin is as follows. The maximum ratio of adducin and spectrin is one adducin per two spectrin tetramers in the adducin-spectrin-actin complexes (Fig. 1D; Fig. 2), consistent with the idea that adducin can interact with more than one spectrin molecule. Electron microscopy of spectrin-actin-adducin complexes provides some examples of two spectrin molecules attached to an actin filament at the same site as adducin⁸. Finally, as discussed below, calmodulin inhibits adducin-dependent binding of spectrin but not binding of adducin to spectrin-actin complexes. Thus, the ability of adducin to promote binding of spectrin may result from a mechanism other than the formation of ternary complexes with spectrin and actin.

Calmodulin, in the presence of micromolar concentrations of free Ca²⁺, inhibits the adducin-dependent binding of spectrin to actin filaments (Fig. 3a). This inhibition is specific for the adducin-dependent binding of spectrin to actin as equivalent concentrations of calmodulin and Ca²⁺ have no effect on binding of spectrin to actin in the absence of adducin, or on binding of adducin to spectrin-actin complexes (Fig. 3b, c). The Ca²⁺ dependence for the inhibitory effect of calmodulin is highly cooperative with respect to Ca²⁺, as is typical for calmodulin-modulated activities⁹, with maximal inhibition at 1 μ M free Ca²⁺ (Fig. 4b).

Adducin is the most probable target for calmodulin in these assays. Calmodulin (in micromolar Ca²⁺) binds adducin with a $K_d \sim 0.2 \mu$ M⁴ and also binds to erythrocyte spectrin, although with a much lower affinity of 6-20 μ M¹⁰⁻¹². Half-maximal effects of calmodulin occur at $\sim 0.9 \mu$ M calmodulin (Fig. 4a), considerably less than the amount required for half-maximal binding to spectrin. Calmodulin is present in erythrocytes at an average concentration of $\sim 3 \mu$ M^{13,14}, it is therefore likely that, under physiological conditions, most of the spectrin sites for cal-

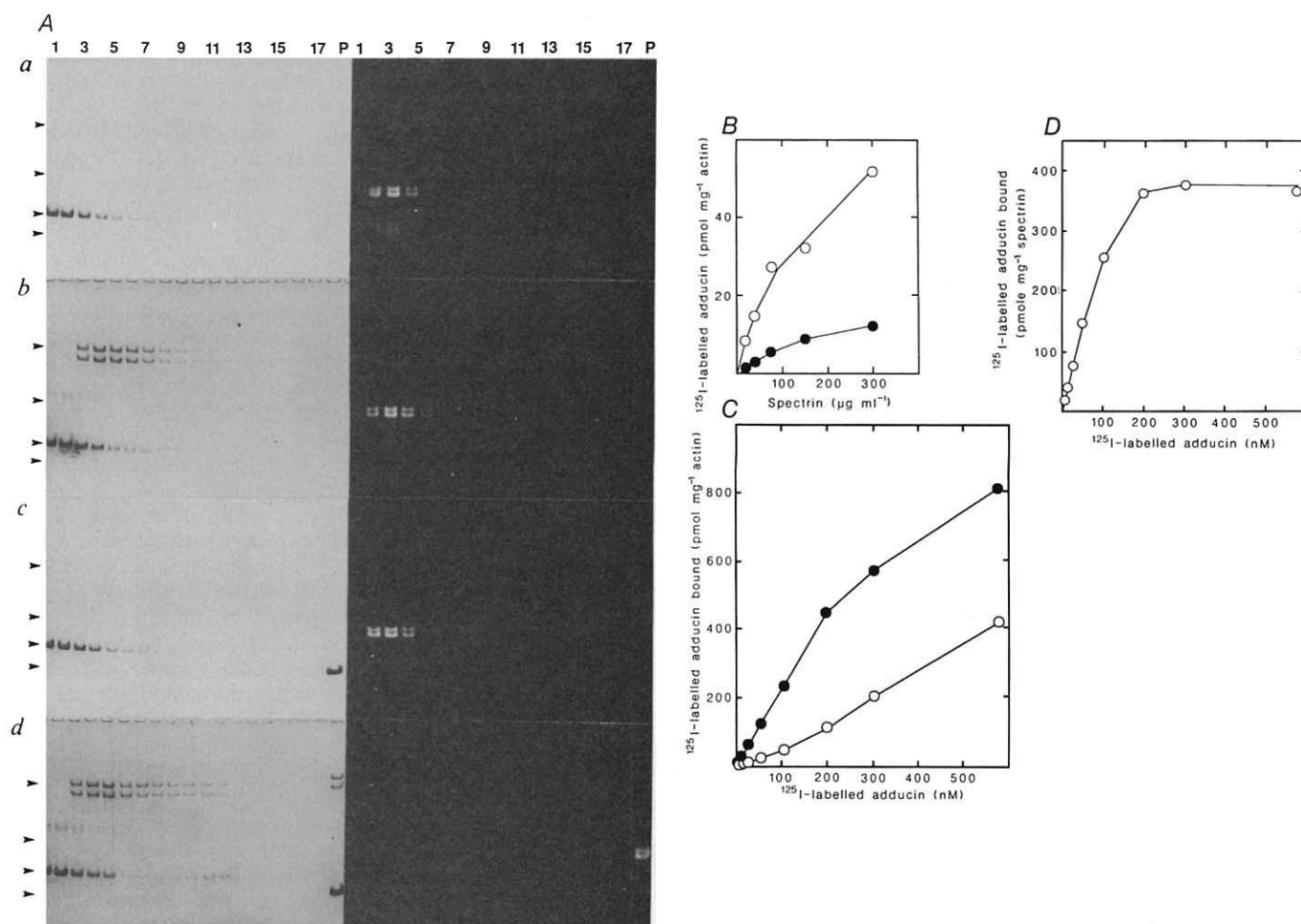


Fig. 1 ^{125}I -labelled adducin binds with high affinity to spectrin-actin complexes but poorly to either spectrin or actin alone. **A**, Coomassie blue (left) and autoradiography (right). **B**, Spectrin-dependent cosedimentation of labelled adducin with F-actin with pig (○) or human (●) erythrocyte spectrin. **C**, labelled adducin cosedimentation with F-actin with (●) and without (○) pig erythrocyte spectrin. **D**, Spectrin-dependent adducin binding to spectrin-actin complexes. **A**, ^{125}I -labelled adducin (100 nM; 9,663 c.p.m. pmol^{-1}) was incubated at 24 °C for 1 h with no additions (**a**); 1.3 μM erythrocyte spectrin (**b**); 3.5 μM F-actin (**c**); or 1.3 μM spectrin and 3.5 μM F-actin (**d**) in 10 mM HEPES, pH 7.3, 30 mM KCl, 50 mM NaCl, 4 mM MgCl_2 , 0.25 mM Na EGTA, 1 mg ml^{-1} bovine serum albumin (BSA), 0.05% (v/v) Tween 20 and 1 mM ATP (Buffer B). At the end of the incubation, 100 μl of each reaction was layered onto 5-ml 10–30% sucrose gradients in buffer B (minus BSA). Gradients centrifuged at 45,000 r.p.m. in a SW 50.1 rotor for 5 h and 0.2-ml fractions collected, analysed by SDS-PAGE, stained and autoradiographed to detect the adducin subunits ($M_r = 103,000$ and 97,000). Arrowheads, M_r markers for spectrin α -subunit, $M_r = 260,000$; adducin α -subunit, $M_r = 103,000$; BSA, $M_r = 68,000$; actin, $M_r = 43,000$. Adducin is not seen in Coomassie blue gels. The lower M_r polypeptides are proteolytic breakdown products of the adducin subunits. **B**, Labelled adducin (32 nM, 12,580 c.p.m. pmol^{-1}) incubated with 3 μM F-actin and 18–300 $\mu\text{g ml}^{-1}$ pig or human erythrocyte spectrin (same conditions as **A**). Reaction mixtures (100 μl) were layered onto 75- μl sucrose barriers (10% sucrose dissolved in Buffer B); sedimented at 42,000 r.p.m. for 1 h at 4 °C in a Beckman 42.2 rotor; supernatant removed and the F-actin pellet analysed for ^{125}I . Binding data are averages of duplicate determinations. **C**, Labelled adducin (6.5–600 nM) binding to 3 μM F-actin was measured $\pm 1.3 \mu\text{M}$ spectrin dimer (see **B**). Labelled adducin (9,663 c.p.m. pmol^{-1}) was added to 3.5 μM prepolymerized F-actin $\pm 1.3 \mu\text{M}$ spectrin dimer in 10 mM HEPES, pH 7.3, 30 mM KCl, 50 mM NaCl, 4 mM MgCl_2 , 0.25 mM Na-EGTA, 1 mM NaN_3 , 1 mM ATP, 1 mM dithiothreitol and 0.05% (v/v) Tween 20 in 100- μl volumes and incubated 1 h at 24 °C. The assay mixtures were then layered onto 75- μl sucrose barriers (10% in the buffers described above) and F-actin-spectrin complexes bound with labelled adducin pelleted by centrifugation (see above). **D**, Bound, labelled adducin was determined from the data in **C** by subtracting the binding in actin from binding in spectrin and actin. Bound spectrin was calculated from the total amount of spectrin, for each data point, by multiplying it with the ratio between the spectrin contained in the supernatant and the spectrin in the pellet fractions. Spectrin ratios were determined by SDS electrophoresis followed by elution of dye from Coomassie blue-stained regions of the gel containing spectrin subunits and measurement of absorbance at 550 nm²⁹.

Methods. G-actin was prepared as described³⁰ followed by gel filtration on Sephacryl S-200, suspended to a final concentration of 2.5 mg ml^{-1} and polymerized to F-actin by the addition of KCl to 30 mM, MgCl_2 to 4 mM and ATP to 1 mM final concentrations. Under these conditions, >90% of the actin was sedimentable. Human and pig erythrocyte spectrin purified as described³¹. Adducin was iodinated with Bolton-Hunter reagent³¹; SDS-PAGE was performed as described³².

modulin are not occupied. The difference in K_d for calmodulin binding to adducin (0.2 μM) and the concentration of calmodulin (0.9 μM) required for half-maximal inhibition of adducin-dependent binding of spectrin to actin may reflect a different affinity of adducin for calmodulin depending on whether adducin is free in solution or if adducin is associated with spectrin-actin complexes.

Our study provides the first evidence for a functional role for adducin in the erythrocyte and demonstrates that: (1) adducin forms a stable complex with spectrin and actin filaments; (2) adducin promotes binding of additional spectrin molecules to actin; and (3) calmodulin at physiological concentrations of Ca^{2+} inhibits binding of spectrin to adducin-spectrin-actin com-

plexes by binding to adducin. We suggest that adducin participates in an ordered assembly pathway with the following steps: (1) spectrin and actin associate to form spectrin-actin complexes; (2) adducin binds to preassembled spectrin-actin complexes; (3) spectrin binds to adducin-spectrin-actin complexes; and (4) calmodulin, in the presence of Ca^{2+} , binds to adducin, causing dissociation of the second spectrin. Confirmation of this mechanism requires more detailed quantitative analysis and measurement of rates of association and dissociation of spectrin and adducin.

The ability of adducin to promote binding of spectrin to actin filaments is reminiscent of the activity of protein 4.1^{15,16}. Adducin and protein 4.1 differ in several important respects,

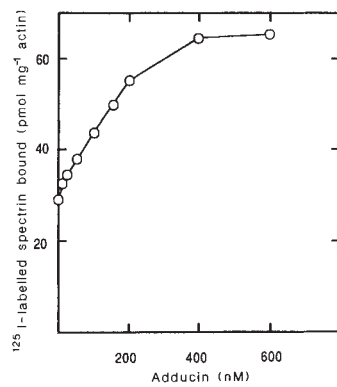


Fig. 2. Adducin stimulates binding of spectrin to F-actin filaments. Binding of ^{125}I -labelled spectrin (100 nM) to 3 μM actin filaments in the presence of various concentrations of adducin. The values for total bound spectrin were corrected for nonspecific binding by subtracting the amount of spectrin that pelleted in the absence of actin. Each point represents the average of duplicate determinations. Binding studies performed as outlined in Fig. 1. Erythrocyte spectrin was iodinated (3,800 c.p.m. pmol^{-1}) with Na^{125}I in the presence of chloramine- T^{33} .

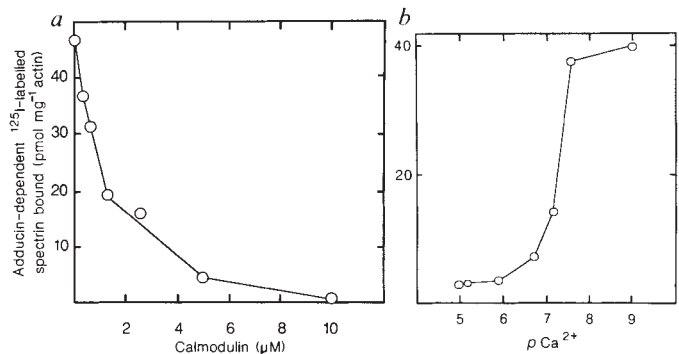


Fig. 4 Calmodulin inhibition of adducin-dependent binding of spectrin to actin is both Ca^{2+} - and dose-dependent. *a*, effect of increasing concentrations of calmodulin on the adducin-dependent binding of ^{125}I -labelled spectrin to F-actin in 10 μM Ca^{2+} . Adducin was present at 200 nM and ^{125}I -labelled spectrin was 400 nM with other conditions as described in *a*. Bound protein represents the fraction of the total bound adducin-dependent ^{125}I -labelled spectrin. *b*, Calcium dependence of the calmodulin inhibition of adducin activity. Free Ca^{2+} (1 nM–10 μM) was incubated with 200 nM adducin, 200 nM ^{125}I -labelled spectrin (9,700 c.p.m. pmol^{-1}), 3 μM F-actin and 10 μM calmodulin. Binding was measured under the conditions described in *a* and shows only the adducin-dependent binding of ^{125}I -labelled spectrin to F-actin. Ca^{2+} concentration expressed as $-\log$ free Ca^{2+} in solution.

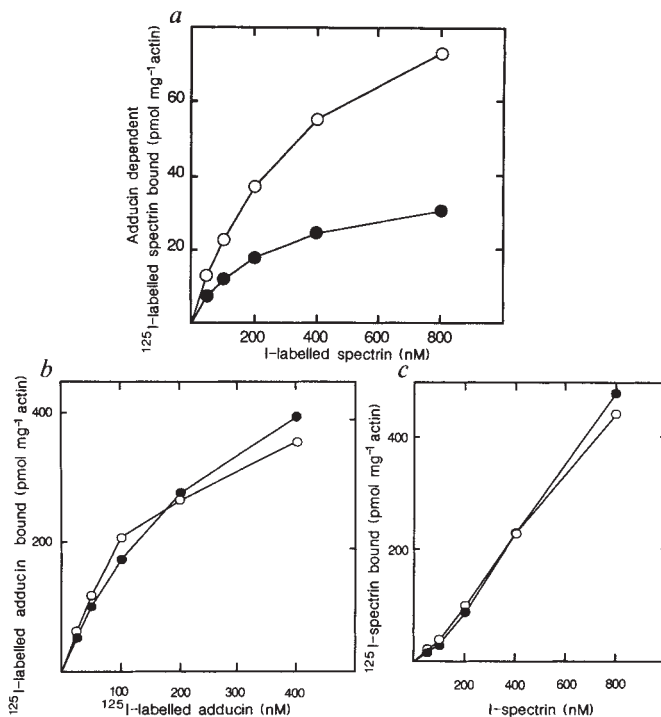


Fig. 3 Calmodulin blocks adducin-dependent binding of spectrin to F-actin. *a*, adducin-dependent binding of labelled spectrin to F-actin as a function of labelled spectrin concentration when incubated with 4 μM free Ca^{2+} in the absence (○) or presence (●) of 10 μM calmodulin. ^{125}I -labelled spectrin (9,666 c.p.m. pmol^{-1}) was incubated with 200 nM adducin and 3 μM F-actin for 90 min at 24 °C. With the exception of the modifications listed above, the rest of the experiment was performed according to *b*. *b*, Calmodulin has no effect on the binding of adducin to spectrin-actin complexes. Labelled adducin was added to 0.5 μM spectrin dimer, 3 μM F-actin and 4 μM Ca^{2+} in the absence (●) or presence (○) of 10 μM calmodulin. *c*, calmodulin has no effect on the binding of labelled spectrin to F-actin. Labelled spectrin (9,700 c.p.m. pmol^{-1}) was added to 3 μM F-actin with 4 μM Ca^{2+} in the absence (●) or presence (○) of 10 μM calmodulin. All other binding conditions and procedures as in *a*. Data represent the mean of duplicate determinations with an average error of <6%. *a*–*c*, Calmodulin purified from bovine brain³⁴; free Ca^{2+} concentrations calculated from $\text{Ca}^{2+}/\text{Na-EGTA}$ ratios and corrected for the presence of KCl and ATP³⁵.

however. Protein 4.1, in contrast to adducin, binds well to spectrin in the absence of actin^{17–19}, and promotes association of spectrin with actin by stabilizing spectrin-actin complexes¹⁵, with an average stoichiometry of spectrin to protein 4.1 that is 1:1 or less^{17,18} and a capacity that is limited only by the total amount of actin in solution.

Adducin is probably involved in biogenesis of spectrin-actin linkages in the erythrocyte membrane skeleton. Protein 4.1 is synthesized late in erythroid development^{20,21}, possibly after assembly of spectrin on the membrane. It is tempting to speculate that adducin mediates early assembly of spectrin-actin complexes in an initially flexible and highly Ca^{2+} -sensitive membrane skeleton that is later stabilized by protein 4.1 as the erythrocyte matures. The proposed ability of adducin to recruit spectrin onto adducin-spectrin-actin ternary complexes may be important for assembly of the spectrin-actin complexes containing 5–7 spectrin molecules gathered about a central actin core which have been visualized in erythrocyte membrane skeletons^{22–25}. This scheme is supported by the fact that erythrocytes completely lacking protein 4.1 can still assemble a normal amount of spectrin and actin on their membranes²⁶.

It is important to evaluate the physiological role of adducin and Ca^{2+} in regulating the membrane properties of erythrocytes. Adducin binds to actin, although with low affinity; perhaps together with other proteins it helps to regulate other properties of actin such as filament length. The consequences of phosphorylation of adducin by protein kinase C or cyclic AMP-dependent protein kinase⁵ have yet to be examined. Immunoreactive forms of adducin are present in brain and other tissues⁴ and have been isolated from brain (V.B. and K.G., unpublished data). Tissue forms of adducin could provide a mechanism for regulation of the distribution of spectrin by Ca^{2+} and calmodulin, and thus have a role in Ca^{2+} -mediated redistribution of spectrin in receptor capping²⁷ and in secretion²⁸.

Supported by NIH, with a Research Career Development Award and by a postdoctoral fellowship.

Received 27 January; accepted 29 May 1987.

1. Marchesi, V. T. *A. Rev. Cell Biol.* **1**, 531–561 (1985).
2. Bennett, V. A. *Rev. Biochem.* **54**, 273–304 (1985).
3. Goodman, S. R. and Zagon, I. S. *Am. J. Physiol.* **250**, 347–360 (1986).

4. Gardner, K. & Bennett, V. *J. biol. Chem.* **260**, 1339-1348 (1986).
5. Ling, E., Gardner, K. & Bennett, V. *J. biol. Chem.* **261**, 13875-13878 (1986).
6. Brenner, S. & Korn, E. *J. biol. Chem.* **254**, 8620-8627 (1979).
7. Bennett, V., Davis J. & Fowler, V. E. *Nature* **299**, 126-131 (1982).
8. Gardner, K. & Bennett, V. *J. cell biochem.* (in the press).
9. Cheung, W. Y. *Science* **207**, 19-27 (1980).
10. Burns, N. & Gratzner, W. *Biochemistry* **24**, 3070-3074 (1985).
11. Husain, A., Howlett, G. J. & Sawyer, W. H. *Biochem. biophys. Res. Commun.* **122**, 1194-1200 (1984).
12. Sears, D. E., Marchesi, V. T. & Morrow, J. S. *Biochem. biophys. Acta* **870**, 432-442 (1986).
13. Agre, P., Gardner, K. & Bennett, V. *J. biol. Chem.* **258**, 6258-6265 (1983).
14. Jarrett, H. & Penniston J. T. *J. biol. Chem.* **253**, 4676-4682 (1978).
15. Ohanian, V. *et al. Biochemistry* **23**, 4416-4420 (1984).
16. Ungewickell, E., Bennett, P., Calvert, R., Ohanian, V. & Gratzner, W. *Nature* **280**, 811-814 (1979).
17. Tyler, J. M., Reinhardt, B. N. & Branton, D. *J. biol. Chem.* **255**, 7034-7039 (1980).
18. Cohen, C. M. & Foley, S. F. *Biochemistry* **23**, 6091-6098 (1984).
19. Eder, P. S., Soong, C. J. & Tao, M. *Biochemistry* **25**, 1764-1770 (1986).
20. Staufenbiel, M. & Lazarides, E. *J. Cell Biol.* **102**, 1157-1163 (1986).
21. Conboy, J., Mohandas, N., Tchernia, G. & Kan, Y. W. *N. Engl. J. Med.* **315**, 680-685 (1986).
22. Beaven, G. H. *et al. Eur. J. Cell Biol.* **36**, 299-306 (1985).
23. Byers, T. J. & Branton, D. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6153-6157 (1985).
24. Shen, B. W., Josephs, R. & Steck, T. L. *J. Cell Biol.* **102**, 997-1006 (1986).
25. Liu, S.-C., Derick, L. H. & Palek, J. *J. Cell Biol.* **104**, 527-536 (1986).
26. Takakuwa, Y., Tchernia, G., Rossi, M., Benabadi, M. & Mohandas, N. *J. clin. Invest.* **78**, 80-85 (1986).
27. Bourguignon, L. Y., Suchard, S. J., Nagpal, M. L. & Glenney, J. R. Jr *J. Cell Biol.* **101**, 477-487 (1985).
28. Perrin, D. & Aunis, D. *Nature* **318**, 585-591 (1985).
29. Fowler, V. M. & Bennett, V. *J. biol. Chem.* **259**, 5978-5989 (1982).
30. Pardee, J. D. & Spudich, J. A. *Meth. Cell Biol.* **24**, 271-289 (1982).
31. Bennett, V. *Meth. Enzym.* **96**, 313-324 (1983).
32. Fairbanks, G., Steck, T. & Wallach, D. *Biochemistry* **10**, 2606-2617 (1971).
33. Hunter, W. & Greenwood, F. *Nature* **194**, 495-496 (1962).
34. Gopalakrishna, R. & Anderson, W. *Biochem. biophys. Res. Commun.* **104**, 830-836 (1982).
35. Fabiato, A. & Fabiato, F. *J. Physiol., Lond.* **75**, 463-505 (1979).

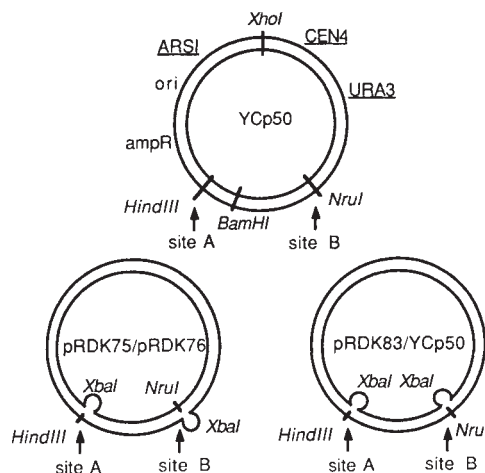


Fig. 1 Heteroduplex plasmid DNAs used in transformation experiments. Genes from pBR322 are in roman type; yeast genes are underlined, and restriction sites are in italic. Plasmid YCp50 (ref. 12) contains the *ColE1 ori*, and the *amp^R* and *tet^R* genes of pBR322, and the *URA3* gene of *S. cerevisiae*, and the selected marker used in yeast transformation. In addition, YCp50 carries the yeast *CEN4* and *ARS1* sequences for stable maintenance of the plasmid. Plasmid pRDK75 was constructed by disruption of the *HindIII* site of YCp50 with an *XbaI* linker, pRDK76 by disruption of the *NruI* site of YCp50 with an *XbaI* linker and pRDK83 was constructed by ligation of the appropriate *BamHI*-*XhoI* fragments obtained from pRDK75 and pRDK76 (ref. 9). Heteroduplexes were constructed and purified from homoduplex DNA by a denaturation and reannealing method¹². Heteroduplex preparations are equimolar mixtures of the two types of plasmid heteroduplexes which are produced by reannealing two different DNAs; only one of the two types of heteroduplex plasmid is shown. The two mismatched sites carried by the substrates are designated A and B; site A corresponds to the *HindIII* site of YCp50 and site B corresponds to the *NruI* site of YCp50. The lengths of the mismatched insertion loops are 12 bp (site A) and 8 bp (site B). The structure of correction products was determined by restriction mapping of plasmid DNA from individual transformed clones⁹. Correction at a given site can convert the mismatch to one of two possible restriction sites. If correction does not occur, replication of heteroduplex DNA results in two distinct daughter plasmids with different restriction sites. The different products can then be distinguished by restriction mapping as previously described⁹.

The role of heteroduplex correction in gene conversion in *Saccharomyces cerevisiae*

Douglas K. Bishop*†, Marsha S. Williamson‡, Seymour Fogel‡ & Richard D. Kolodner*§

* Laboratory of Molecular Genetics, Dana Farber Cancer Institute, Boston, Massachusetts 02115, USA

† Department of Cellular and Developmental Biology, Harvard University, Cambridge, Massachusetts 02138, USA

‡ Department of Genetics, University of California, Berkeley, California 94720, USA

§ Department of Biological Chemistry, Harvard Medical School, Boston, Massachusetts 02115, USA

Two different models have been proposed to explain the relative frequencies of the non-mendelian allelic segregations which are detected by tetrad analysis after meiosis in fungi¹⁻⁷. The first model maintains that 6:2 type tetrads result from correction of heteroduplexes containing mismatched sites and 5:3 type tetrads result from failure to correct mismatched sites. The second model suggests that 6:2 segregations result from the filling-in of double-strand gaps using information obtained from both strands of a homologous duplex. In this model 5:3 type tetrads result if the allele is included in the heteroduplex regions flanking the gap and the resulting mismatched nucleotides are not corrected⁷. We have studied the correction of heteroduplex plasmid DNA in *pms1* mutant strains of *Saccharomyces cerevisiae*, which are known to exhibit higher frequencies of 5:3 type tetrads and lower frequencies of 6:2 tetrads than wild-type strains. Our results suggest that the *pms1* mutation causes a defect in mismatch correction, supporting the hypothesis that meiotic gene conversion in wild-type yeast cells often results from the correction of heteroduplex DNA.

Yeast *pms1/pms1* homozygous diploid strains show increased frequencies of post-meiotic segregation and reduced frequencies of 6:2 type segregation relative to *pms1/PMS1* or *PMS1/PMS1* cells⁸. The *pms1* mutations are also known to cause a mitotic

mutator phenotype and to increase the frequency of meiotic recombination between closely linked markers: all these effects can be explained if *pms1* mutations cause a defect in mismatch correction. To determine if *PMS1*, the wild-type allele at this locus, is involved in mismatch correction in mitotic cells, the effect of the *pms1-2* mutation on correction of heteroduplex plasmids was tested. Details of the method used to construct heteroduplex plasmid DNA and the restriction mapping method used to characterize the correction products obtained from individual transformants have been reported previously⁹. Both of the heteroduplex substrates carry two small (12- and 8-base pair (bp)) insertion-deletion mismatches separated by 943 bp; the two mismatched sites are referred to as site A and site B, respectively (Fig. 1). The two substrates differ only in that pRDK75/pRDK76 carries the two insertion loops on opposite strands and pRDK83/YCp50 carries the two insertion loops on the same strand (Fig. 1). The two different heteroduplex conformations were used to distinguish between effects on the efficiency of co-correction and effects on the parity of correction (as will be discussed below). Two closely related haploid strains were transformed; MW3022-9D is wild-type (*PMS1*) and MW3022-4B carries the *pms1-2* mutation. The data are presented in

† To whom correspondence should be addressed.

Table 1 Analysis of individual colonies obtained by transformation with heteroduplex DNA

Transformant classes observed *										
pRDK75/pRDK76	I	II	III	IV	V	VI	VII	VIII	IX	N
										60
wild type †	38%	42%	3%	10%	3%	3%	0%	0%	1%	
<u>pms1-2</u> ‡	15%	15%	0%	19%	2%	27%	0%	5%	15%	58
pRDK83/YCp50	I	II	III	IV	V	VI	VII	VIII	IX	N
										103
wild-type	6%	11%	38%	34%	6%	0%	0%	0%	6%	
<u>pms1-2</u>	2%	1%	2%	30%	2%	36%	2%	10%	16%	53

Key:  corrected by insertion  corrected by deletion  or  not corrected

The DNA from individual transformants was purified and digested with a mixture of *Xho*I and *Xba*I and a mixture of *Xho*I, *Hind*III and *Nru*I. Digested samples were run on 0.8% agarose gels and transferred to nylon filters. Filters were hybridized to ³²P-pBR322 to detect predicted plasmid DNA fragments. This procedure (see ref. 9 for details) allowed unambiguous assignment of transformants into one of the nine predicted correction classes. The diagrams under each Roman numeral represent inferred structures of plasmid DNA after repair and before the first round of replication. Plasmids are represented in linear form for convenience; all plasmid DNAs detected were covalently-closed circular monomers.

* Transformants in classes I-IV are interpreted as resulting from correction at both mismatched sites and plasmid DNAs from these transformants have a unique restriction map. Transformants in classes V-IX result from failure to repair at one or both mismatched sites and plasmid DNA from these transformants yields restriction endonuclease cleavage patterns characteristic of a mixture of two distinct plasmids.

† Recipient strain MW3022-9Dα (*trp1Δ, ura3-52, leu1, his7-2, ade2-1*).

‡ Recipient strain MW3022-4Bα (*trp1Δ, ura3-52, leu2, his7-2, ade2-1, pms1-2*).

Table 1, which also indicates how classes of transformants with each plasmid are designated.

Most wild-type transformants (70-80%) contained plasmids with a configuration of restriction sites identical to one of the two plasmids used to construct the heteroduplex (pRDK75/pRDK76 classes I and II, pRDK83/YCp50 classes III and IV). These 'parental' transformants can be accounted for by correction of both mismatched sites; the transformants in these classes could be generated by two independent correction reactions on the same strand or a single co-correction event². Roughly 15% of the transformants had a recombinant configuration of restriction sites which most probably resulted from independent correction of both mismatched sites on opposite strands (pRDK75/pRDK76 classes III and IV, pRDK83/YCp50 classes I and II). Some (6%) of the transformants contained a mixture of parental and recombinant plasmid DNA, which is consistent with correction of only one of the two mismatched sites followed by replication (classes V to VIII for both substrates). All of the single-site correction transformants resulted from repair at site B (classes V and VI). Failure to correct both sites (class IX) occurred in <6% of the transformants. The results obtained with this strain agree with earlier data obtained by transforming another wild-type strain (ref. 9; D.K.B. and R.D.K., unpublished observations).

The data with *pms1-2* differ from those with wild type in three conspicuous ways (Tables 1 and 2). First, the overall efficiency of correction of both pRDK75/pRDK76 and pRDK83/YCp50 was reduced. This is reflected by a decrease in recovery of classes I-IV which are generated by correction of both mismatched sites, and by an increase in classes V-IX which are generated by failure of correction at one or both mismatched sites. In *pms1* cells, as in wild-type cells, site B was corrected more efficiently than site A, so that there were more

class V and class VI transformants than class VII and VIII transformants (Tables 1 and 2). Second, plasmid DNAs recovered from *pms1* transformants seem to result from correction events which favour deletion of the small loops, whereas analysis of the wild-type transformants suggested that the heteroduplex plasmid DNA was corrected with parity. This is most simply illustrated by comparing the proportion of products produced by one or two deletion events (classes IV, VI and VIII) to the proportion of products formed by one or two

Table 2 Calculated correction parameters

		Efficiency of correction*		Disparity in favour of deletion‡
		site A	site B	
pRDK75/pRDK76	wild-type	93%	99%	1.1
	<i>pms1-2</i>	56%	80%	3.1
pRDK83/YCp50	wild-type	88%	94%	0.8
	<i>pms1-2</i>	46%	72%	9.0

Values were calculated from the data in Table 1. Statistically significant differences ($P \leq 0.01$) referred to in the text were determined by Fisher exact tests¹³.

* The efficiency of correction at site A is given by $[(I + II + III + IV + VII + VIII) / N] \times 100$, where roman numerals refer to the number of transformants in each class and N is the total number of transformants examined. The efficiency of correction at site B is given by $[(I + II + III + IV + V + VI) / N] \times 100$.

‡ Disparity of correction in favour of deletion is given by (sites repaired in favour of deletion)/(sites repaired in favour of insertion) = $[I + II + 2(IV) + VI + VIII] / [I + II + 2(III) + V + VII]$, using the same symbols.

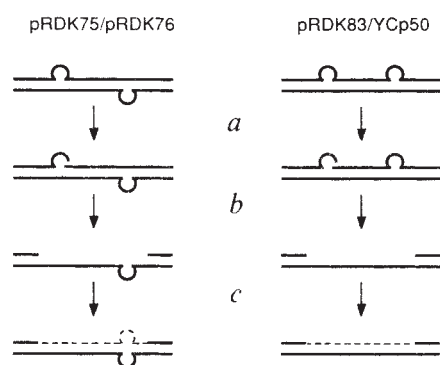


Fig. 2 Co-repair model to account for the distribution of single-genotype transformants obtained after transformation of *pms1* cells with pRDK75/pRDK76 and pRDK83/YCp50. *a*, Initiation of correction occurs at or near a single-strand insertion loop (in wild-type cells correction events start on either strand). *b*, Excision tract elongation past the adjacent mismatched site. *c*, DNA synthesis across excised region (newly synthesized DNA is represented by dotted lines). See Table 1 for a description of the symbols used. The product of the co-correction reaction can be either pRDK75 (class I) or pRDK76 (class II) when pRDK75/pRDK76 is the substrate. Because of the disparity of the initiation reaction, only YCp50 (class IV) is produced from co-correction of pRDK83/YCp50.

insertion events (classes III, V and VII). Third, the disparity of correction following transformation of the *pms1* strain with pRDK83/YCp50 is three times greater than that observed with pRDK75/pRDK76 (Table 1). The difference in the degree of disparity results from the substantial number of pRDK75/pRDK76 transformants in classes I and II. Class I and II transformants are generated by a deletion event and a coincident fill-in (insertion) event at the adjacent site.

The data can be accommodated by a model which suggests that the *PMS1* gene product is required for productive correction events that are initiated on the non-insertion strand opposite single-strand insertion loops. In a *pms1* mutant strain, most correction events initiate on the insertion strand at or near the single-strand loop (Fig. 2). The exact function of *PMS1* in correction is unclear at present. It is possible that it is required either to make incisions opposite single-strand insertion loops, or for the initiation of excision and/or resynthesis opposite insertion loops. The model explains the decrease in correction efficiency and the disparity of correction observed in the mutant strain. The model also explains the difference in the degree of disparity observed by transforming the *pms1* strain with pRDK75/pRDK76 and pRDK83/YCp50 if co-correction occurs in *pms1* cells. pRDK75/pRDK76 insertion events (classes I and II, Table 1) could result from co-correction when excision-resynthesis initiates at an insertion loop and proceeds past the adjacent mismatch (Fig. 2). Co-correction would not lead to insertion events with the pRDK83/YCp50 heteroduplex. We have assumed that single-genotype transformants (classes I–IV) result from either independent correction or co-correction of both mismatched sites. However, single-genotype transformants might be generated by random strand loss before correction is complete¹⁰. We have argued that the observed excess of wild-type pRDK75/pRDK76 transformants in classes I and II compared with classes III and IV results from co-correction of the two sites rather than from strand loss. The observation (Table 1) that single-genotype transformants containing a parental configuration of restriction sites are reduced in the *pms1* strain supports the idea that the excess of parental-genotype transformants is due to co-correction rather than strand loss.

Our results are consistent with genetic studies on the effect of *pms1* mutations on meiotic recombination and support the hypothesis that strains carrying *pms1* mutations are deficient in

correction of heteroduplex DNA both during mitosis and meiosis⁸. If heteroduplex correction deficiency is the only defect conferred by the *pms1*-2 mutation, then at least some meiotic 6:2 type gene conversion events in wild-type cells result from correction of heteroduplex DNA. This raises questions about the proportion, if any, of meiotic 6:2 type segregations that result from inclusion of a mutant allele within the double-strand-gap region as postulated by the double-strand-break repair model⁷. It is possible that *pms1* mutations are pleiotropic and that mutant strains are deficient in both repair of heteroduplex DNA and double-strand gaps. Recent results suggest that *pms1* mutants are not grossly deficient in double-strand-gap repair, although further experiments are required to determine if the *pms1* mutation alters the average length of heteroduplex DNA tracts formed during double-strand-break repair events (A. Nicolas, M.S.W., J. Szostak and S.F., unpublished results). Note that, although it is possible that the *pms1* mutation affects both heteroduplex formation and heteroduplex correction, the observed effect on correction is sufficient to explain the meiotic phenotype of *pms1* strains.

Received 3 March; accepted 22 May 1987.

1. Fogel, S., Mortimer, R. K., Lusnak, K. & Tavares, V. *Cold Spring Harb. Symp. quant. Biol.* **43**, 1325–1341 (1979).
2. Fogel, S., Mortimer, R. K. & Lusnak, K. in *The Molecular Biology of the Yeast Saccharomyces* Vol. 1 (eds Strathern, J. N., Jones, E. W. & Broach, J. R.) 289–339 (Cold Spring Harbor Laboratory, New York, 1982).
3. Fogel, S., Mortimer, R. K. & Lusnak, K. in *Molecular and Experimental Perspectives* (eds Spender, J. F. T., Spencer, D. M. & Smith, A. R. W.) 65–107 (Springer, New York, 1983).
4. Radding, C. M. *A. Rev. Biochem.* **47**, 847–880 (1978).
5. Golin, J. E. & Esposito, M. S. *Molec. gen. Genet.* **183**, 252–263 (1981).
6. Meselson, M. S. & Radding, C. M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 358–361 (1975).
7. Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J. & Stahl, F. W. *Cell* **33**, 25–35 (1983).
8. Williamson, M. S., Game, J. C. & Fogel, S. *Genetics* **110**, 609–646 (1985).
9. Bishop, D. B. & Kolodner, R. D. *Molec. cell. Biol.* **6**, 3401–3409 (1986).
10. Wagner, R. & Meselson, M. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4135–4139 (1976).
11. Stinchcomb, D. T., Mann, C. & Davis, R. W. *J. molec. Biol.* **158**, 157–179 (1982).
12. Fishel, R. A. & Kolodner, R. *J. molec. Biol.* **188**, 147–157 (1986).
13. Fleiss, J. L. *Statistical Methods for Rates and Proportions* (Wiley, New York, 1981).

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem

☐ Other (please specify) _____

☐ I am interested in sending my order by mail.

☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____

Title _____

Institution/Company _____

Department _____

Address _____

City _____ State _____ Zip _____

Phone (____) _____

UMI Article Clearinghouse

Mail to: University Microfilms International
 300 North Zeeb Road, Box 91 Ann Arbor, MI 48106